

THE IMPACT OF EVALUATION IN COMMUNITY COLLEGE
FACULTY EFFORT AND EFFECTIVENESS

By

ELIZABETH CAROL COLLINS

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN
PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

1988

ACKNOWLEDGMENTS

My thanks and appreciation go to Dr. J. L. Bettenberger, Dr. James Longstrech, and Dr. John Jones, my committee members, for their guidance and encouragement, and especially for their patience in working with an out-of-town distressed student who was attempting to balance the demands of writing a dissertation with the demands of a new position as a community college administrator. There is no way adequately to say thank you to Dr. Bettenberger, who taught each of us the meaning of teacher, scholar, and friend, and for us, helped bridge the gap between educational theory and practice.

Special thanks go to three special friends, Mike, Roger, and Bob, who encouraged me constantly, always believed I would make it, never allowed me to quit or even to slow down, and were there to share the good times and the bad.

An acknowledgment goes to my colleagues at St. Peter's-Hug Junior College who were willing to accept far less than 100% effort from me for two years and who often did more than their share so that I could balance work and study. Very special thanks go to Tom Lang and Pat Cappi, who so expertly and unselfishly accepted my responsibilities while I was away and to my boss and friend, Dr. Ray Church,

for encouraging me to make my degree my priority and for teaching me so much about the art of educational administration.

And last, and most importantly, my love and thanks go to my family for their support. My Mom and Dad were willing to put up with my absences, demands, and irritability and always gave me their unquestioning love and encouragement. My deepest love and appreciation go to my son Brett, who made the greatest sacrifices and did without his mother at a time when he needed her. My promise goes to him to always be there for him in the future and to share with him the joys and benefits of education.

TABLE OF CONTENTS

ACKNOWLEDGMENTS	13
ABSTRACT	19
Chapter	
I. INTRODUCTION	1
Statement of the Problem	11
Limitations and Delimitations	12
Justification	13
Assumptions	14
Definitions	15
Procedures	17
Sampling Design	17
Collection of Data	17
Data Analysis	18
Summary	22
II. REVIEW OF LITERATURE	23
The Evolution of Faculty Evaluation	24
Motivation	27
Expectancy Motivation Theory	28
Path-Goal Theory	34
Intrinsic Motivation	38
The Theory of Evaluation and the Exercise	
of Authority	47
Frequency of Evaluation	47
Research in Universities	48
Contingency Theories	51
Summary	58
III. METHODS AND PROCEDURES	60
The Problem	63
Questionnaire Design	65
Sample Selection	66
Data Collection Procedures	68
Analysis of the Data	78
Summary	79

IV. RESULTS	74
Description of Project Sample and Response	74
Responses Related to Perceived Frequency of Evaluation	77
Responses Related to Leverage	79
The Relationship Between Perceived Frequency of Evaluation and Leverage	81
Results from Classifying Teams Along the Active-passive Continuum	87
Results from Analyzing the Effort and Effectiveness Components of Leverage Separately	89
Results from Analyzing Perceived Frequency of Evaluation and Perceived Leverage in a Single Stage of the Evaluation Model	93
V. SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS	95
Summary	95
Project Rationale and Design	96
Design	99
Analysis	100
Conclusions and Implications	109
Recommendations	109
APPENDICES	
A. QUESTIONNAIRE	113
B. LETTERS	121
REFERENCES	125
BIOGRAPHICAL SKETCH	139

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

THE IMPACT OF EVALUATION ON COMMUNITY COLLEGE
FACULTY EFFORT AND EFFECTIVENESS

By

Elizabeth Carol Collins

August, 1988

Chairman: Dr. J. L. Wittenborgues
Major Department: Educational Leadership

The relationship between the perceptions of community college faculty relating to the frequency of evaluation activity and their perceptions of their own effort and effectiveness was examined in this study. Drawing on the theory of evaluation and the theories of authority presented by Durkheim and Weber in their 1878 book, Evaluation and the Exercise of Authority, and on the work of Macriello presented in a report to the Spencer Foundation in 1981, titled The Impact of the Evaluation of Teaching on Teacher Effort and Effectiveness, the purpose of the study was to investigate whether there was a statistically significant relationship between perceived frequency of evaluation and perceived leverage, a term defined as the ratio of perceived

effectiveness to perceived effort. Additional purposes were to investigate whether this relationship was influenced by the ways in which community college faculty conceived of their tasks in terms of predictability, efficacy, and clarity, whether perceived frequency of consultation was more strongly related to perceived effectiveness or to perceived effort, and whether perceived frequency of consultation in any one step of the consultation model was more strongly related to leverage. These questions were investigated for three six teaching tasks: classroom teaching, curriculum development, professional development, student advising, college service, and assistance lectures.

A sample of 500 faculty members was drawn from the 28 community colleges in Florida using a probability proportionate to size sampling strategy. A mail-out questionnaire was used to collect data.

The relationship between perceived frequency of consultation and perceived leverage was found to be significant and moderately positive for the tasks of curriculum development and college service and significant and moderately negative for the task of classroom teaching. The strongest relationships were found for tasks that faculty perceived as low in predictability, clarity, and efficacy. The steps of the consultation model with the strongest relationships between

perceived frequency of evaluation and leverage were affecting, sampling, and planning for improvement.

Based on these data, community college administrators might consider increases in certain kinds of evaluation for certain fields as a way to increase the leverage of faculty.

CHAPTER I

INTRODUCTION

After several decades of rapid expansion and growing resources, community colleges are now facing two decades of declining enrollments, declining financial resources, and, at the same time, an increasing public emphasis upon accountability and the improvement of quality. This ground-swell across the nation for accountability has come from taxpayers, institutional trustees, financial donors, and students, all of whom are pressuring colleges and universities to establish the cost-effectiveness of each department and the performance of each professor (Seldin, 1984).

At the same time, academic departments are increasingly "tunered-in," resulting in a less-mobile and aging college faculty and the "entrenchment" of whole faculties, which James Bean (1983) called "economic and organizational incarceration" (p. 23). E. E. Walton (1980), in a study of high commitment work systems, anticipated that the results of this entrenchment will be a shift from the high motivation of faculty stemming from a strong sense of self-direction and accomplishment to apathy or even apathy. This faculty may become hostile, both privately and publicly, towards educational institutions which offer few challenges

and a two-percent future [Oberholzer, 1984]. With these threats at the desktop, the question is frequently asked, "What can be done to increase faculty satisfaction, motivation, and productivity?"

Attempts to find answers to this question have resulted in an increased interest in faculty evaluation as a possible solution. However, interest in faculty evaluation is not new. In 1942, in The Academic Man, Louis Hilsen wrote:

It is an exaggeration to say that the most critical problem confronted in the social organization of any college or university is the proper evaluation of faculty services, and giving due recognition through impartial assignment of honor. (p. 212)

Hillar (1972) described faculty evaluation as one of the most complex and expensive issues facing the academic world. Kane (1971) noted:

Cost-effectiveness of operations will be more carefully examined. If this is not done internally, it will be done externally by the new agencies working for legislators and governors. (p. 20)

The need for improved faculty evaluation systems still exists today and, if anything, is increasing exponentially as enrollment declines, costs climb, and lifespans increase (Giddin, 1988). Many regard evaluation as an important component in increasing faculty motivation, satisfaction, and productivity. Others disagree and regard it as a procedure that is of limited value but that is inevitable due to fiscal problems and pressures from legislators, students, and the public (Hillar, 1972).

Print (1985) pointed out that "evaluation is an inherent element in our organized attempt to achieve a goal" (p. 187). Kerr (1978) described evaluation as the key component of all MBO (management by objectives) systems of organizational planning and controlling. Others, however, contended that teaching is an art and far too complex to be reduced to a set of questions or too subtle to be observed perceptively by outsiders (Higbet, 1990). Wilson (1988), commenting on the subjective nature of faculty evaluation, argued that "there is no valid or reliable instrument for measuring teacher effectiveness, nor will there ever be" (p. 18).

Faced with the opposing forces of a pressing need to develop evaluation procedures and a general lack of consensus on methods, procedures, and desirability of evaluating faculty, a plethora of research dealing at length with methods and procedures for assessing faculty performance has been reported in the last two decades (Gentle, 1981; Wilson, 1981). Researchers have investigated, reported, and summarized hundreds of studies on the validity, reliability, sources of evidence, and on other topics related to the mechanics of faculty evaluation.

Evaluation is most often treated as a method of achieving greater control over the educational process, thereby serving as a method for improving the efficiency and effectiveness of schools and colleges. Little attention has been

given to the impact of such evaluations on faculty and thus on students (Karr-Saunders, 1914).

When considering the effects of evaluation, one of the first philosophical issues which arises is the tension between the concepts of academic freedom and academic responsibility. The right to teach, study, and conduct research without interference, a concept implicit in the principle of academic freedom, is often cited as a strong argument against evaluation (Hopie, 1982).

Another issue relating to the effects of evaluation which has gained widespread attention in recent years is the effects of evaluations on the intrinsic motivation of faculty. McGeachie (1982) found in his review of studies of faculty motivation that faculty find their positions satisfying in terms of complexity, autonomy, interpersonal relationships, and responsibility for other persons. Deci (1975) and White (1958) found that intrinsic motivation is based on the innate need to be competent and self-determining.

Opponents of evaluation believe that the use of external evaluations undermines the intrinsic motivation to teach by decreasing autonomy and self-determination and by decreasing opportunities for creativity and innovation (Deci & Ryan, 1980). However, Morgan (1984) indicated that evaluations, properly administered, may actually increase intrinsic motivation.

One of the first comprehensive studies of the effects of evaluation was reported by Bernbach and Scott [1975] in their book, Evaluation and the Burden of Authority, and was the result of a research program at Stanford University on evaluation and authority in a variety of organizations. Based on studies conducted in schools, they suggested that there are benefits to be derived from increasing the frequency of evaluation. They found a positive relationship between frequency of evaluation (formal and informal) and increased teacher satisfaction (Bernbach & Scott, 1975).

The Bernbach and Scott [1975] theory is useful for the investigation of the effects of evaluation on teachers because it is based on an explicit model of the evaluation process. The model originally specified four stages in a complete model for the evaluation of performance.

Assignment is the process of assigning a task to an individual performer. Tasks may be allocated by directive (giving specific instructions as to how to carry out the assignment) or by delegation (assigning responsibility for final results).

Criteria setting is the process of determining which properties of a task should be considered in the assessment of performance, deciding the relative weight of each property, and setting a standard for performance along each dimension.

Sampling is the process of collecting information on the performance of the allocated task according to the criteria for performance that have been set.

Appraisal is the process of bringing together the criteria set for the task with the sample of information collected on that performance to arrive at an evaluation.

Although not explicitly listed as steps in the Borschuch and Scott (1975) model, two additional steps were explicit in the process.

Communicating is the process of communicating the results of evaluation to the faculty member.

Planning for improvement is the process of developing a plan to help faculty members improve behavior and the securing of resources for this plan.

In another part of their research, Borschuch and Scott (1975) focused on the variation of tasks performed by faculty and suggested that different work arrangements and supervisory behaviors may be called for depending upon the task. They developed the idea of "task conceptions" to refer to faculty beliefs concerning the characteristics of their tasks (Borschuch & Scott, 1975).

Borschuch and Scott (1975) suggested in their theory of evaluation and authority that task conceptions will vary along an active-passive dimension. Tasks are variable in the extent to which goals can be clearly specified (clarity), in the extent to which the resistance to performing the task is

predictable (predictability), and in the extent to which means can be developed for achieving the desired outcomes (efficacy). Tasks which are high on clarity, predictability, and efficacy are described as inert tasks, and those that are low on these three dimensions are described as active tasks. Borschbach and Scott (1971) demonstrated that there is a relationship between performers' task conceptions and their preferences for supervisory and work arrangements. They found that when task conceptions are active, performers prefer to receive assignments in a way that allows them to make significant decisions about the procedures for carrying out these tasks, but that when task conceptions are inert, performers prefer to receive directions that provide them with predetermined courses of activity (Borschbach & Scott, 1971).

Warraslin (1984) used a six-step model of the evaluation process, a task-specific focus, and the concept of task conceptions, all part of the Borschbach and Scott (1971) theory of evaluation and the exercise of authority, to examine the effects of evaluation systems on teacher effort and effectiveness. He defined his six-step model of the evaluation process as the four-step Borschbach and Scott (1971) model plus the two additional steps of communicating and planning for improvement, which were implicit in the Borschbach and Scott model. He chose not to follow the lead of earlier evaluation researchers who used either effort or

outcome is missing the object of evaluation on workers. His rationale was that neither effort nor effectiveness was satisfactory when considered alone. A measurement of effort ignores final outcomes of that effort, and a measurement of outcome is open to the criticism that it fails to take into account the differences in task difficulty and the variations in the predictability of accomplishing different tasks and might discourage organizations and individuals from taking on new, riskier tasks (Mottishaw, 1988).

Mottishaw (1984) developed a new concept to take into account both effort and outcomes. He described this new concept called "leverage" as follows:

To overcome the weaknesses of these two approaches we developed a concept that takes into account both effort and outcomes. This new concept, leverage, refers to the relationship between the effort put forth by a subordinate and the outcomes resulting from that effort. Leverage involves the relationship between a worker and his or her work, or more specifically, between the effort exerted by a worker and the outcomes or product of that effort. Leverage can be described as the ratio between outcomes and effort and may range from 1:0 to 0:1. 1:0 represents the case where for 0 units of effort the individual realizes 1 unit of output; 0:1 represents the case where the individual realizes 0 units of output for every 1 unit of effort. Of course, 1:1 represents the case where for every unit of effort the individual realizes 1 unit of outcome. (p. 548)

The concept of leverage is closely tied to Vroom's (1964) expectancy motivation theory. Vroom postulated that the force of motivation is the product of expectancy, valence, and instrumentality. Expectancy is defined as a worker's perception of the relationship between the level of

effort and the level of performance. Herzai, Berman, and Wilson (1988) summarized numerous studies and concluded that the force of motivation is significantly related to job satisfaction and performance. Yessierli (1994) concept of expectancy, which is one part of the force of motivation, is almost identical to Herzstein's (1984) concept of leverage.

The use of the concept of leverage shifts the focus of managerial strategies, including evaluation, from the goal of increasing managerial control over efforts or outputs to enabling subordinates to gain control over their own work (Herzstein, 1984). Hackman and Wageman (1995) described control (experienced responsibility for work outcomes) as one of the three psychological states necessary for maintaining intrinsic motivation. Argyris (1983) stated control as a necessary ingredient for the sustained motivation of mature adults.

In his study of 188 middle school teachers, Herzstein (1984) hypothesized that the more frequent the evaluation of teaching, the greater will be the leverage of teachers up to the point of overly close supervision. His second hypothesis was that the more active (lower on clarity, predictability, and efficacy) the task conception, the greater the impact of evaluation on leverage up to the point of a task being too active for supervisors to make useful contributions. The second hypothesis was based on the rationale that because most tasks are by definition very predictable

and very straightforward, performers are unlikely to find that increased supervision will increase the ratio of their output to their effort. In addition, because active tasks are very unpredictable and very unclear, supervisors may be able to make no contribution to improving the ratios between output and effort.

Metcalfe (1984) did not confirm his first hypothesis because he found a positive relationship between perceived frequency of evaluation and increased leverage at all levels of evaluation activities. There was no decrease in the relationship at high levels of supervision as he had hypothesized. He was not able to test his second hypothesis, because the teachers in the study perceived all their tasks to be highly inert (Metcalfe, 1984).

Are these implications in this far community college administrators striving to design evaluation systems and to implement management strategies to address the pressures and problems discussed earlier? While no generalizations can be made from data collected from middle school teachers, Metcalfe (1984) suggested the possibility that frequency of evaluation is related to faculty leverage (effectiveness/effort) at all educational levels and that the extent of this relationship is dependent upon task conceptions (active vs. inert)....

Statement of the Problem

The purpose of this study was to examine the relationship between the perceptions of community college faculty of the frequency of evaluation activity and their perceptions of their own effectiveness in relationship to their effort on these six teaching tasks: classroom teaching, curriculum development, professional development, college service, student advising, and maintenance activities.

Specifically, the following questions were addressed:

1. For each of the six teaching tasks, was there a statistically significant association ($p < .05$) between perceived frequency of evaluation and perceived effectiveness in relationship to effort [leverage]?

2. Was there a stronger relationship between perceived frequency of evaluation and leverage for tasks that faculty perceived as active, for tasks faculty perceived as moderately active, or for tasks faculty perceived as inert?

3. For each of the six teaching tasks where there was a statistically significant association ($p < .05$) between perceived frequency of evaluation and leverage, was perceived frequency of evaluation related to increases in perceived effectiveness, decreases in perceived effort, or both?

4. In which stages of the evaluation process did perceived frequency of evaluation have the strongest

association with perceived effectiveness in relationship to effort (leverage) for a combination of the six teaching tasks?

Limitations and Delimitations

1. The study was limited to full-time faculty members at five community colleges in Florida, and the results should not be generalized beyond the setting from which the sample was drawn.

2. The study attempted to determine if a relationship existed between perceived frequency of evaluation and perceived effectiveness in relationship to effort. No attempt was made to show a causal relationship.

3. Data on frequency of evaluation, effort and effectiveness, and task conceptions were gathered solely from responses to questionnaire items.

Justification

In a 1973 report of the Assembly of the American Association of Community and Junior Colleges (1974) stated:-

In academic terms, the staff is the college's most significant and largest capital investment. In these hard times, we affirm that it is only good sense that the investment should be helped to appreciate in value and not be allowed to wear itself out or slide into obsolescence by neglect or neglect (p. 128)

It is a well-known fact that 70 to 80% of the operating budgets of community colleges are spent on faculty salaries and fringe benefits (Wilson, 1980)- it seems

reasonable to conclude that if colleges are to address the problems of declining enrollments, declining financial resources, and an increasing public pressure for accountability and increased quality, the solutions must involve improving faculty effectiveness.

As discussed earlier, Borchert and Scott (1973) and Kretzsch (1984) found a relationship between perceived frequency of evaluation and perceived feelings of effectiveness in relationship to effort in studies conducted in elementary schools. If an analysis of the data collected from community college faculty substantiates this hypothesis, then community college administrators interested in enhancing faculty perceptions of effectiveness, and very possibly actual effectiveness, might consider increased formal and informal evaluation as a tool for accomplishing this goal. If the hypothesis is not substantiated, increased evaluation might be considered inappropriate as a method of increasing faculty perceptions of effectiveness in relationship to effort. An analysis of the data collected on task conceptions may also guide administrators as to which of the six teaching tasks, if any, are most likely to be positively affected by increased evaluation.

If no significant relationship is found between perceived frequency of evaluation and leverage, those who believe that evaluation is a threat to the intrinsic motivation of faculty and those who view it as a threat to the

principles of academic freedom will have additional evidence on which to base their argument.

However, if a significant relationship is found between perceived frequency of evaluation and leverage, this research may be used to silence the opponents of evaluation who perceive it as a control mechanism limiting the autonomy and academic freedom of faculty (Patriscio, 1984). The opponents, who have long slowed progress in the area of evaluation, may come to see evaluation as a positive practice in enhancing faculty control over their teaching tasks. This research may be used to put us "properly the longstanding philosophy that college faculty deserve complete autonomy and that anything other than complete autonomy is injurious to the tradition of academic freedom. It may be a step towards silencing the spreading cry of evaluation opponents who believe that evaluation decreases intrinsic motivation which ultimately decreases faculty satisfaction and faculty effectiveness. It may be used to provide a new management philosophy to community college department chairpersons who, by tradition and training, have not dared to "interfere" in the work of faculty.

Assumptions

1. It was assumed that the responses from the participants were honest and could be taken at face value.

2. It was assumed that non-responses were random in nature and therefore did not significantly change the results of the study.

Definitions

Terms used in this study are defined as follows:

Active task refers to a task that a faculty member perceives as less predictable, clearer, and efficient.

College service refers to work done by a faculty member outside of the classroom such as serving on committees and representing the college at civic activities.

Community college refers to a two-year public institution which offers programs and courses limited to the first two years of postsecondary education and includes university parallel programs, occupational education, and continuing education.

Effectiveness refers to the success a faculty member experiences in achieving desired outcomes on a particular task.

Effort refers to how long and how hard a faculty member works on a particular task.

Evaluation refers to any way a supervisor communicates either directly or indirectly to the subordinate as to how well the subordinate is doing on a particular task. It includes everyday compliments and criticisms, nonverbal communication, as well as formal evaluation procedures.

Full-time faculty member refers to a faculty member who is under contract for a full-time teaching load for at least one academic year.

Just task refers to a task that a faculty member perceives as high on predictability, clarity, and efficiency.

Intrinsic motivation refers to motivation which results from an experience that is an end in itself; an activity that is rewarding in itself, not just as a means to an external goal.

Leverage refers to the relationship between effort put forth by a faculty member and the outcomes resulting from that effort (Natraville, 1984).

Maintenance factors refer to the routine responsibilities of a faculty member, such as recordkeeping, office hours, and required reporting.

Perceived frequency of evaluation refers to the perceptions of a community college faculty member as to the frequency of formal and informal evaluation activities by their immediate supervisors.

Task perceptions refer to the ways faculty members view the attributes of their different tasks in terms of predictability, clarity, and efficiency (Gershbach & Scott, 1973).

Teaching tasks refers to different tasks which characterize the work of a community college faculty member. For purposes of this study, these tasks are defined as classroom teaching, curriculum development, student advising,

professional development, college services, and maintenance factors.

Procedures

Sampling Design

The sample used for this study was selected from a target population of all full-time faculty members in the 28 community colleges in Florida. A multistage cluster sampling technique with probability of cluster selection proportional to size was used so that each faculty member had an equal chance of selection. Using this process, 5 of the 28 community colleges were selected and 100 full-time faculty members were randomly selected from each of the 5 colleges. Colleges containing less than the required 100 full-time faculty members were combined so that each combination contained the required number. The total sample size was 500.

Collection of Data

The following steps were followed to obtain the data for the study.

1. A list of the 28 community colleges in Florida and the number of full-time faculty members at each was developed using college catalogs. Colleges with less than 100 faculty members were combined to form one cluster. The list was weighted using probability proportional to size

techniques, and five clusters were selected using computer-generated random numbers.

2. Using a table of random numbers, 100 full-time faculty members were selected from each of the five clusters.

3. A survey questionnaire was constructed containing questions requiring Likert-scale response categories which resulted in summary measures of perceived frequency of evaluation, perceived fairness, and conceptions of the six teaching tasks along the active-inert continuum.

4. The questionnaire was pretested with 25 faculty members who were not selected as part of the sample to test for question clarity. Modifications were made as necessary.

5. The questionnaires, with stamped return envelopes and cover letters, were mailed to the 500 faculty members in the sample.

6. A list of non-responders, plus duplicate questionnaires and return envelopes, was provided to a key administrator at each institution who had previously agreed to serve as a coordinator for the study. Follow-up continued until 81% of the total sample had responded.

Data Analysis

Using the methods of Norusis (1984), data obtained from the survey responses were used to calculate summary measures of perceived frequency of evaluation, perceived

leverage, and conceptions of the six teaching tasks along the active-passive spectrum according to the degree of task clarity, predictability, and efficiency.

For each of the six teaching tasks, faculty were divided into three groups: those reporting no stage of the six-stage evaluation model; with frequent activity, those reporting one stage with frequent activity, and those reporting two or more stages with frequent activity.

Two measures of leverage were obtained from the faculty surveys. The Faculty Assessment of Leverage (FAL) was obtained from an item asking faculty to describe the relationship between their effort and the payoff from that effort in terms of results during the last year. On the basis of the responses, faculty were classified according to those for whom the payoff was less than their effort and those for whom the payoff was equal to or greater than their effort. The FAL measure was obtained for each of the six tasks.

A second measure of leverage, the Composite Assessment of Leverage (CAL), was constructed from two other items on the faculty survey which asked faculty separate questions about their perceptions of their effectiveness and their perceptions of their effort. Faculty were placed into two response groups by taking the ratio of effectiveness to effort. Those whose effectiveness was less than effort were one group, and those whose effectiveness was greater than

effort were a second group. The CAL measure was calculated for each of the six teaching tasks. Unlike the PAL measure, the CAL measure was designed to allow for an analysis of whether perceived frequency of evaluation had a stronger relationship to perceived effort or to perceived effectiveness.

Three items on the questionnaire were used to measure task conceptions for each of the six tasks. For each task, responses were collapsed into a high, medium, or low category reflecting ratings along the achievement dimension. The median response category was used to classify each task as high, medium, or low in terms of perceived position along the achievement continuum.

A summary measure was also developed for each separate stage of the evaluation process to indicate frequency of activity in each stage for a combination of the six tasks. Faculty responses to questions relating to frequency of evaluation were collapsed into high, medium, or low categories for each stage of the evaluation process. A summary measure of perceived leverage for a combination of all tasks was developed by asking faculty to describe their overall effectiveness in relationship to effort.

To answer question number 1 of the problem statement, gamma's were calculated for each of the six teaching tasks using both the CAL and PAL measures of leverage to determine whether there was a significant association between

perceived frequency of evaluation and perceived effectiveness in relationship to effort (leverage) for community college faculty. Gamma is a non-parametric measure of association between two ordinal variables (Coster, 1988; Sie et al., 1975). Each gamma was tested for statistical significance at the .05 level (Dunn & Krutz, 1988).

To answer question number 2 of the problem statement, gammas were calculated using frequency and leverage data combined for all tasks faculty considered high on the active-inert dimension, combined for all tasks faculty considered medium along the active-inert dimension, and combined for all tasks faculty considered inert along the active-inert dimension. Gammas were compared to see if the relationship between perceived frequency of evaluation and perceived leverage was stronger for tasks conceived as being more active.

To answer question number 3 of the problem statement, the separate measures of effort and effectiveness provided by the GLE surveys were used. Gammas describing the association between perceived frequency and perceived effectiveness and between perceived frequency and perceived effort were compared to determine whether perceived effort or perceived effectiveness had a stronger relationship to perceived frequency for each task where there was a significant association between perceived frequency of evaluation and perceived leverage.

To answer question number 4 of the problem statement, graphs describing the relationship between perceived frequency of evaluation activity in each stage of the evaluation process and perceived leverage were compared to determine whether increasing evaluation in a particular stage might be more effective than increasing it in another stage.

Summary

Background information and justification for a study of the relationship between frequency of evaluation and perceived effectiveness in relationship to effort for a variety of teaching tasks which vary along an achievement continuum have been presented in Chapter I. An overview of the literature, including the evolution of faculty evaluation, expectancy motivation theory, intrinsic motivation, the theory of evaluation and the exercise of authority, and the relationship of other contingency models of leadership to the concept of task conceptions will be presented in Chapter II. The development and testing of the survey instrument, the sampling design, and the procedures for collecting and analyzing the data will be described in Chapter III. The research findings will be presented and discussed in Chapter IV, and a summary of the study, conclusions, analysis, recommendations, and implications will be presented in Chapter V.

CHAPTER 11

REVIEW OF RELATED LITERATURE

In reviewing the literature in the area of faculty evaluation, the dominant theme that emerges is that faculty evaluation is in a state of growth and continuing change. It is a science that is still in its infancy and will experience such dramatic changes in the next decade that it may be hardly recognizable in its present form by the year 2000 (Feldin, 1994). As Arthur Chickering (1986) noted, "The days of benign neglect and just-of-the-paste evaluation are fast disappearing" (p. 81).

In this chapter, the literature relating to the evolution of faculty evaluation systems up to the current time is reviewed with emphasis on the new trends which are affecting faculty evaluation in colleges. The effects of evaluation on motivation are discussed. This includes a review of expectancy motivation theory as it relates to the concept of leverage and a review of the effects of evaluation on intrinsic motivation. The Bandura and Mott (1975) theory of evaluation and the exercise of authority is reviewed in terms of their research relating to education and in terms of their concept of task conceptualization in relationship to other contingency theories.

The Evolution of Faculty Evaluation

The changes in faculty evaluation closely paralleled the changing conditions in higher education. Since 1870, greater care and systematization have been evident in evaluation practices and procedures. Economic and political pressures have increased the demand for accountability. Changes in faculty governance in the form of increased unionization and shared control with nonacademic administrators have increased the demands from faculty for more viable and equitable evaluation systems. The expansion of the size of institutions and the number of faculty in each institution has made the traditional personal and collegial evaluation procedures obsolete (Baker, 1976).

Gale (1984) likened the changes that have taken place in higher education that relate to evaluation as growth from "monastery to industry" (p. 100). Until the eighteenth century, students hired and dismissed their own teachers and paid greater fees to the more esteemed teachers (Gallin, 1880; Mendels, 1897). Practices similar to this are still seen today in some informal disciplines such as music and in teacher plans for educational finance (Cople, 1982).

Even as the traditional organization of colleges and universities that we know today began to emerge in the eighteenth and nineteenth centuries, Gale (1984) said they still resembled a monastery in terms of faculty evaluation. He described the faculty as

a group of dedicated, otherworldly, underpaid, and somewhat socially deviant males--accepted as knowledge--ideally bonded together to pursue individual, isolated, and often esoteric searches for understanding and enlightenment. They engaged in scholarship, and they taught the young.

Each professor's commitment was to his own learning and, through it, to his discipline. His institution was a vehicle that permitted him to pursue his interests, to grow throughout a lifetime according to his own pattern. That growth was almost totally self-governed and self-monitored. Like a good Victorian child, he knew what was expected of him; he had internalized the rules and systems of control. (p. 100)

In this environment, accountability and evaluation depended on careful selection of competent, self-monitoring professors. Peers made decisions when necessary to determine whether a candidate was acceptable for employment or continued membership in the group (Bordia, 1988).

Changes which occurred in higher education after World War II began to change this idyllic environment. Most of the changes related to growth. As universities increased in size, both the student body and the faculty became more heterogeneous. As older, less passive students began enrolling in colleges, they demanded more control over their education (Doyle, 1943). As organizational structures became more complex, more decisions, including those involving evaluation, were taken away from faculty (Borde, 1970). Increases in financing by outside agencies further imploded on the traditional decision-making processes in colleges and universities (Muller, 1978).

The emergence of the utilitarian philosophies of the aucties brought another surge in growth in colleges and universities and spurred the beginning of community college systems in many states. With this change in philosophy came the consumer movement and the tremendous pressures of student unrest (Miller, 1979).

The result of these changes was that higher education became an industry and the life of the professor became a job (Kain, 1984). Vigorous activity was evident in the area of evaluation as colleges and universities struggled to replace the controls of old traditions and values with new systems and procedures.

With the invention, the new problem of declining financial resources came to higher education, and, as might be expected, the evaluation impetus grew stronger. During the sixties, entire books began to appear on instructional evaluation (Mellum, 1971; Osmer, 1981; Carter, Greenough, & Hengen, 1971; Kullik & McNeachie, 1979; Miller, 1982, Miller, 1975). The science of faculty evaluation had changed from small amounts of fragmented research into a new discipline (Boyer, 1981).

The majority of evaluation research up until the eighties was centered on the mechanics of evaluation: student ratings vs. classroom observation vs. self-evaluation, the validity and reliability of various evaluation methods, the use of student juries to evaluate

faculty performance, and the formation of, sensitive uses of evaluation information. However, in the eighties, higher education was faced with yet a new problem. Dramatically declining enrollments in many institutions resulted in little opportunity to hire new faculty and left colleges with an aging faculty, hired during the expansion of the sixties, faced with little ability, reduced funds for research, and less opportunity to choose the courses they taught (Hess, 1982).

A 1971 address by E. F. Christ to the National Association of State Universities and Land Grant Colleges described the ideas which were beginning to emerge even as early as the seventies.

The problem of how to bring new faculty members in is becoming the problem of how to counsel and coach out. Those happy recruitment parties at the scholarly conventions are being replaced by dreary technical meetings on the substantive foundations of early retirement. (Miller, 1975, p. 3)

With these changes, faculty evaluation has quickly taken on a new dimension. Attention has turned from the mechanics of evaluation, which occupied so much of the time of researchers and theorists in the sixties and seventies, to the new problem of how to use evaluation as a way to motivate faculty and to increase faculty satisfaction in a profession that is not as promising as it once was.

Motivation

The level of theoretical activity involving motivation in work organizations has increased during the last few

years. Administrators have identified understanding the individual in the work organization as their most difficult problem (Mishel, Boffstein, & Wilson, 1980). Latham and Ericsson (1978) attributed this attention to motivation to the increase in human problems such as absenteeism, mis-aiming, and lack of commitment to work. Dierks and Porter (1978) suggested that there is a new recognition that behavioral factors are of critical importance if organizations are to attract and retain dependable and creative performers and that human resources are long-term assets. They further suggested that pressures from increased competition and pressures from external agencies have made it necessary to seek new mechanisms to increase effectiveness and efficiency.

Several hypotheses, theories, and approaches to organizational effectiveness have resulted from the study of organizational behavior. What makes some organizations work more efficiently or become more productive than others? More specifically, what factors are involved in making individuals in an organization more efficient or more productive than others within or outside of the organization?

Two topics relating to motivation in organizations are closely linked to the study of the effects of evaluation on the effort and effectiveness of community college faculty. The first, expectancy motivation theory, has been a dominant approach to research on motivation for many years (Campbell

& Pritchard, 1974). The second, the study of intrinsic motivation, has only gained in importance in recent years.

Expectancy-Motivations Theory

Mudrack (1981) said that

expectancy theory focuses on the cognitive processes through which individual effort is energized, directed, and sustained over time. Expectancy theory attempts to explain how individuals make decisions between different alternatives or activities (such as spending time preparing a lecture or on research) and decisions about how much effort to invest in any particular activity. (p. 88)

Numerous theoretical variations on expectancy theory exist (Georgopoulos, Mahoney, & Jones, 1982; Lawler, 1972; Staw, 1976; Vroom, 1964), but most share common theoretical components. House (1971) stated that

the forces on an individual to engage in a specific behavior is a function of (1) his expectations that the behavior will result in a specific outcome and (2) the sum of the valences, that is personal utilities or satisfactions, that he derives from the outcome and that an individual chooses behaviors he engages in on the basis of (1) the valence he perceives to be associated with the outcomes of the behavior under consideration, and (2) his subjective estimate of the probability that his behavior will indeed result in the outcome. (p. 322)

Perhaps more clearly stated, House and Mitchell (1974) postulated that an individual's attitude and behavior can be predicted from: (1) the degree to which a behavior is seen as leading to various outcomes (expectancy), and (2) the evaluating of these outcomes called (valences). If people are satisfied with a job, it is because they think it leads

to things that are highly valued, and they work hard if they believe that effort leads to these highly valued things.

Gilbreath and Cummings (1997) developed a model to explore the interrelationships of expectancy, valence, and instrumentality. Expectancy is defined as the relationship between behavior and performance levels and is high, for example, if a faculty member believes that it is highly probable that intensive efforts to increase preparation will lead to increased student achievement and positive student attitudes (Mischel, DeFrance, & Milroy, 1988).

Valence is defined as the importance an individual attaches to a reward that is received from working in an organization and may include recognition, salary, personal satisfaction, and other rewards (Kewley, 1982).

Instrumentality is defined as the perceived probability that a reward will follow a given performance (Kewley, 1982). For example, instrumentality is high if a faculty member perceives that high student achievement or positive student attitudes are likely to result in increased financial or other rewards (Mischel, DeFrance, & Milroy, 1988).

The basic assumption of expectancy motivation theory is that individuals make conscious decisions as to how to behave and that these decisions are influenced by evaluating probabilities and the importance of certain outcomes. In expectancy theory, the focus of motivation is the prospect of

expectancy, valence, and instrumentality (Miskel, Gellman, & Wilson, 1980). The relationship between these components and examples of each is presented in an adaptation of the Gellman and Cummings (1987) model in Figure 1.

Hersman and Schweb (1978) and Mitchell (1974) in comprehensive reviews of research on expectancy motivation theory found that the force of motivation is positively correlated with job satisfaction and performance. However, in most studies, large amounts of the variance remained unexplained, and in a few studies, contradictory results were reported (Miskel, Gellman, & Wilson, 1980).

Although expectancy, instrumentality, and valence have been recognized in most studies as interrelated components in the determination of the force of motivation, much greater attention has been given to the instrumentality and valence components than to the expectancy component. Neither Gellman and Scott (1979) nor Bettrillo (1984) mentioned expectancy motivation theory, but the constructs used in their research are clearly related to expectancy motivation theory.

Betrillo's concept of leverage, which is expressed as the relationship between perceived effort and perceived effectiveness, is closely related to expectancy, which has been defined as the subjective probability that effort will lead to the desired performance outcome (Wooday, 1981). Since expectancy is equal in terms of importance to

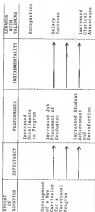


Figure 1. An Illustration of Expectancy Motivation Theory.

instrumentality and valence in the determination of the force of motivation, the importance of leverage in determining performance is apparent:

One of the criticisms of Vatterliere's (1984) research is that he relied on perceptions rather than actual data on effectiveness and effort. If the numerous studies (Bassens & Schenk, 1973; Mitchell, 1974) in which it was found that expectancy, as one part of the expectancy motivation formula, is related to the force of motivation was correct, then the perception of effort in relationship to effectiveness may be a determinant of performance, regardless of whether these perceptions are accurate reflections of actual effort and effectiveness. Therefore, any supervisory behavior, including evaluation, that is related to increases in the perception of leverage may be important in determining organizational effectiveness.

Gallweith and Cummings (1987) also described another part of the expectancy motivation theory that is related to the study of the effects of frequency of evaluation on faculty leverage. One of the examples of valence that Gallweith and Cummings (1987) used was recognition of effort. Evaluation is clearly a mechanism by which a department chairman can provide recognition. However, the relationship between evaluation and the components of the expectancy motivation theory becomes stronger if considered

in relationship to path-goal theory which is an extension of the expectancy activation theory.

Path-Goal Theory

Path-goal theory is in large part an expansion of expectancy theory concepts into the leadership domain (House & Mitchell, 1974). According to House and Mitchell (1974), is the path-goal theory of leadership, leaders are effective because of their impact on subordinates' motivation. They suggested that a leader is motivating or satisfying to the degree that the leader's behavior increases subordinates' goal attainment and clarifies the paths to these goals.

One of the first groups to introduce these theories under the title of path-goal theory was Georgopoulos, Mahoney, and Jones (1967). They accepted the notion that individuals have certain goals and that by achieving these goals, needs of the individual are met. Further, this goal-directed behavior is a result of rational decision making on the individual's part. They postulated that motivation depends on the particular needs of the individual as reflected in the goals towards which he is moving and his perception regarding the relative usefulness of productive behavior as an instrumentalism or as a path to attainment of these goals. They also identified freedom from barriers to follow the desired path as another factor that affects productive behavior (Georgopoulos, Mahoney, & Jones, 1967).

Although this first path-goal theory formulation contained many of the same ideas as the expectancy motivation theory, the concept of freedom from barriers to follow a chosen path was added and thus set the stage for later studies which added the new variable of leadership behavior as a determinant of the force of motivation.

Two other individuals are acknowledged as the major, and most contemporary, contributors to the path-goal theory of leadership. Martin Evans (1974) noted that the motivational process starts with the supervisor behaving in a certain way. Evans defined two types of behavior. The first, labeled consideration, involves indications of trust, respect, concern for personal needs, two-way communication, and participation in decision making. The second type of behavior is labeled initiation of structure and focuses on organizing and defining work, establishing the roles of the subordinates, and striving for the desired performance.

These two forms of behavior affect the subordinates by influencing the perception of path-goal instrumentality (the extent to which following a certain path or behaving in a certain way is seen as leading to goal attainment). However, as Evans (1974) pointed out, motivation is only one contributor to actual job performance. Environmental factors, the nature of the task, and individual abilities are also important factors.

This position agrees with that of Cummings and Schwab [1978] who suggested that there are three basic factors that influence performance in organizations: ability, motivation, and climate.

In the path-goal approach, leadership behavior is considered an independent variable having a direct effect on specific psychological states of the subordinate (House & Bennis, 1974). House and Bennis [1974] listed intrinsic job satisfaction, satisfaction with extrinsic rewards, the expectancy that effort leads to effective performance, and the expectancy that performance leads to rewards as some of the psychological states of subordinates that may be affected by supervisory behavior.

Bennis [1974] extended his theory to include three types of moderator variables in the relationship between supervisor behavior and subordinates' perceived path-goal instrumentalities. These moderator variables are the degree of upward influence that a supervisor has in the organization, the amount of role conflict the subordinate is experiencing, and the extent to which the subordinate is predominantly interest or external in control. He found that the greater the upward influence of the supervisor, the more likely he could influence the expectancies of subordinates. Also, the higher the role conflict, the less likely a supervisor is to be able to influence expectancies, and the more externally

controlled the subordinate is, the stronger the relationship between leader behavior and expectations.

The second major contributor to path-goal theory, E. J. House (1971), presented a view of motivation as a combination of intrinsic and extrinsic valences and path instrumentalities and then showed how these might be influenced by leader behavior. House expressed the leader-subordinate relationship in four general propositions:

1. The motivational function of the leader consists of increasing personal pay-offs to subordinates for work-goal attainment and making the path to these pay-offs easier to travel by clarifying it, reducing blocks, and increasing opportunities for personal satisfaction.
2. The leader's instrumental behavior will have a positive motivational effect to the extent that it reduces role ambiguity or makes possible the exercise of externally imposed controls. Internally imposed controls are effective only to the extent that they allow the allocation of positive valences contingent upon desirable rewards that are controlled by the leader.
3. Where leader attempts to clarify path-goal relationships which are redundant with existing conditions, these attempts will result in increased externally imposed controls that may be seen by subordinates as redundant, resulting in a possible increase in performance, but in a decrease in satisfaction.
4. Instrumental behavior of the leader will increase performance only to the extent that such behavior increases the positive valences associated with goal-directed behavior. (p. 513)

Path-goal theory has been extended in more recent formulations to include an additional step that attempts to

specify types of leadership behavior that are appropriate in different situations (Follett, House, & Kerr, 1971).

House and Bieschke (1974) demonstrated that task structure generally determined whether leader initiation and consideration contributed to subordinate satisfaction, performance expectations, and role clarity. The strongest findings concerned the relationship between instrumental (directive) leadership and subordinate satisfaction. It was found that instrumental leadership is negatively correlated with satisfaction under highly structured tasks and that this relationship increased to a positive correlation as task structure decreased (House & Bieschke, 1974).

Although authors such as House, Mitchell, Bieschke, and Evans strongly supported the path-goal theory, Scitloggi and Sims (1974) and Gross (1969) showed mixed results in their studies.

Many of the findings from path-goal theory research are closely tied to the Barnhouse and Scott (1975) and Macriello (1988) research on the relationship between frequency of evaluation and job satisfaction and performance. In addition, the Barnhouse and Scott (1975) concept of tasks that vary along an active-inert dimension is clearly a variation on the path-goal theory proposition that task structure is a moderating variable affecting the types of leadership behavior which are effective in a particular situation. It is apparent that evaluation is one method that supervisors

might use to help subordinates see effort as leading to effective performance and performance as a path leading to the attainment of personal goals.

One factor which has caused much of the controversy over the path-goal and expectancy motivation theories is the failure to include intrinsic rewards as outcomes having positive valence. Galbraith and Cummings (1987) have observed that the traditional model can be modified to include the idea of intrinsic rewards. Bach (1978) suggested that a model combining the two types of expected outcomes, extrinsic and intrinsic, has greater predictive and explanatory value than a model considering only extrinsic factors.

Intrinsic Motivation

In discussing the motivation of faculty, Katz (1977) stated:

Present views of the problems of poor teaching in higher education are limited by the relative lack of recognition of the importance of intrinsic motivation in instructional staff. Unless faculty perceives the teaching experience as a continuing source of profound satisfaction or life-satisfaction arising out of the fulfillment of deep-seated human needs--they will rarely have the sustained role commitment that is necessary for creativity and excellence in performance. (p. 141)

Katzbach (1981) noted that as financial pressures increase and the demands for accountability grow, the natural administrative reaction is to tighten control of faculty. Colleges develop more formal and systematic

evaluation system, courses offered because of faculty interests are eliminated, variations in teaching load are reduced, and irregular work hours are discouraged (Linn, 1982).

McBeath (1982) concluded that while there are some positive effects to intrajob evaluation, there are also some unintended negative effects. One negative effect mentioned is that evaluation increases anxiety and that faculty who become anxious under the threats of evaluation are likely to be less creative, more rigid, less effective in solving problems, and to use less effective methods of teaching and learning. Greenfield and Weissenberg (1986) found a natural strain in human relations as the result of evaluation.

The second negative effect mentioned by McBeath (1982) is that while evaluation may reduce poor teaching resulting from inadequate preparation and lack of attention to teaching responsibilities, it is also likely to result in less innovation, less emphasis on long-term educational outcomes, and less enthusiasm and commitment to teaching. Lusk (1989) found that "a rating scale becomes a shaping device as neither how supportive the supervisor, the principal, or the system" (p. 18). Collaenderhøj (1974) argued that this kind of lack of enjoyment in teaching causes the teaching activity to lose its primary value for both the faculty member and the student, because when a teacher teaches only for extrinsic reasons (pay, high evaluations, etc.), he

curriculum has own time and conveys the message to the student that learning is only a means to an end.

In describing intrinsic motivation, Hackman and Olden (1980) noted that there are 40 internal states valued and pleasurable psychological states:

1. Experienced Meaningfulness of the Work--the degree to which the individual experiences the job as one which is genuinely meaningful. Does it demand a variety of skills and talents. Does it require completion of a whole and identifiable piece of work, and does it have a substantial impact on the lives of other people?

2. Experienced Responsibility for Work Outcomes--the degree to which the individual feels personally accountable and responsible for the results of the work he or she does. Does the work provide autonomy in task performance?

3. Knowledge of Results--the degree to which the individual knows and understands on a continuous basis how effectively he or she is performing on the job. Are there built-in formal and informal feedback mechanisms? (p. 147)

Now (1977) maintained that whether or not faculty achieve these valued inner states is largely dependent on conditions of work which are subject to design by administrators.

In terms of the three intrinsically rewarding states proposed by Hackman and Olden (1980), are faculty likely to see them as valuable and are they likely to see these states as forthcoming from the act of teaching itself? Many researchers (Moss, 1967; Cohen, 1973; Lewis, 1976) have concluded that most faculty question whether what they do in preparation for teaching will result in good teaching

Because of the many variables that affect the path between faculty effort and student gain, many faculty doubt that what they do does make a difference in the lives of their students (Cohen, 1971).

The second of the psychological states proposed by Becker and Giddens (1988)--experienced responsibility for work outcomes--is also a problem for faculty. Because American higher education is founded on a strong tradition of academic freedom, faculty have generally not been monitored or held accountable for their teaching (Boyer, 1981). Although there have been many cautions against possible negative effects of evaluation, Fain and Andrews (1988) noted that complete autonomy without some constraints, whether bureaucratic or normative, results in lowered productivity. Levin (1974) noted that the fact that research productivity, other service activities, or longevity are given much greater weight than teaching in decisions on salary and promotions reinforces the perception of faculty members that they have little accountability for their teaching efforts. Bass (1981) suggested that this lack of accountability plus the difficulties in determining whether teaching efforts are productive leads to a low probability that faculty will perceive a relationship between their performance and the desired psychological state of experienced responsibility for work outcomes.

Because of the same factors that make it improbable that faculty will experience the first two of the intrinsically rewarding states delineated by Hackman and Wageman [1995], it is also unlikely that they will experience the third state--knowledge of results. Ross [1977] noted that it is widely agreed that faculty are not treated as the art and science of teaching and thus often do not know how to find clues that their efforts are productive. Hagstron [1986] wrote that because of the specialized nature of subject matter and because of the norms of individualism and academic freedom, college instruction is a lonely occupation. This lack of feedback and isolationism is a denial of the needs for belonging and for status and esteem.

Ross [1977] summarized the disbarriers to the achievement of intrinsically rewarding states in this statement:

Suffice it to say that the ambiguity of both formal and informal norms and evaluation criteria on most campuses is a source of anxiety for many faculty. They are denied access to the pleasures to be obtained from achievement of the idiosyncratic objectives, forced into roles that provide little intrinsic reward, experience the anxiety of being controlled rather than controlling, and find both their needs for affiliation and for play and novelty stifled by inhibitions about potential violations of norms of sharing. (p. 138)

Holl and Ransawer [1982] called intrinsic motivation a "want-free" input that universities cannot afford to waste; Ross [1977] criticized colleges and universities for emphasizing teaching improvement programs rather than teacher improvement programs. He examined the typical faculty

solutions to the common difficulties in teaching and observed a pattern of defensiveness, self-protectionism, and a reduction in the importance attached to teaching.

Deci (1984) summarized a plethora of research studies on the conditions which are likely to diminish and enhance intrinsic motivation. In some of this research, rewards of all kinds were found to decrease intrinsic motivation. In other research, competition (Deci & Ryan, 1980), surveillance (Lippert & Green, 1975), and evaluation of performance (Smith, 1974) were also found to decrease intrinsic motivation.

Deci and Ryan (1988) discussed a theory called cognitive evaluation theory and, based on this theory, stated that rewards, surveillance, competition, and evaluation change the perceived locus of causality from internal to external and leave a person feeling less self-determining. When an activity is engaged in to get a reward or to comply with a constraint the activity becomes instrumental and is no longer engaged in for its own sake (Deci & Ryan, 1988).

Lockerman et al. (1978) found that when people are given a choice about which task to engage in or how to proceed on a task, they are more intrinsically motivated.

Deci and Ryan (1988) also found that positive feedback leads to greater intrinsic motivation and that negative and negative feedback tend to decrease motivation. This is tied to the second part of cognitive evaluation theory which Deci

and Ryan (1980) interpreted as meaning that intrinsic motivation is tied to perceived competence.

Other researchers have begun to find evidence that contradicts the idea that rewards decrease intrinsic motivation. Deci (1971) found that rewards have both a controlling and an informational component. He hypothesized that to the extent an activity is seen as a means to an end, in other words is instrumental in receiving a reward, the task will be seen as less inherently interesting and will be less likely to be engaged in if rewards are absent. If, on the other hand, rewards convey information on a person's competence on a task, intrinsic liking for the task will increase and the task will be more likely to be engaged in if rewards are absent. This research has been supported in studies by Smith and Ross (1978) and Ryan (1982).

Deci and Ryan (1981), commenting on the conditions which motivate faculty, wrote:

Times are becoming more difficult in colleges and universities, and the pressure appears to be building. Thus, it is increasingly imperative that administrators work to create the conditions that promote rather than drive intrinsic motivation. The keys to creating these conditions seem to be administrators' orientations toward autonomy and structures that impose neither than control. If conditions are created that bolster teachers' feelings of competence and self-determination, teachers' motivation to teach will take care of itself. (p. 18)

Widely Cribbenworth: (1978) studied intrinsically motivated individuals and found that regardless of the activity involved, intrinsically motivated persons regularly

experience great enjoyment and deep involvement with the task at hand, which he called "flow."

Csikszentmihalyi (1981) listed those things that can be done in planning the teaching environment to create "flow" experiences for faculty:

1. The activity should be structured so that the actor can increase or decrease the level of challenges being faced in order to match exactly his or her skills with the requirements for action.
2. It should be easy to isolate the activity at least at the perceptual level from other stimuli--external or internal--that might interfere with involvement in it.
3. There should be clear criteria for performance; one should be able to evaluate how well or poorly one is doing at any time.
4. The activity should provide concrete feedback to the actor so that one can tell how well one is meeting the standards of performance.
5. The activity ought to have a broad range of challenges--possibly several qualitatively different ranges of challenges--so that the actor may obtain increasingly nuanced information about different aspects of the self. (p. 25)

Although it is evident that there are no conclusive answers as to exactly what determines the intrinsic motivation of faculty, it is clear that evaluation is an important factor to be considered in determining how to change the teaching environment so that faculty experience teaching as an activity that is rewarding in itself.

The Theory of Evaluation and the Exercise of Authority

Bornbush and Scott (1955) developed the theory of evaluation and the exercise of authority after a decade-long research program in a wide variety of organizations. There are 18 propositions with numerous theorems and corollaries relating to power and authority in organizations in the complete theory. The parts of the theory and research that relate either to higher education or to the effects of frequency of evaluation are considered in this review.

Frequency of Evaluation

Using workers in five organizations, Bornbush and Scott (1955) tested the hypothesis that frequency of evaluation is negatively correlated with the organizational status of the respondent. They found no relationship between reported frequency of evaluation and position in the organization. Teachers included in this study reported a median frequency of evaluation of " seldom " for all four teaching tasks.

In this same study, they found that frequency of evaluation was positively correlated with satisfaction with the evaluation process (Bornbush & Scott, 1955).

In another study, they again tested the relationship between frequency of evaluation and satisfaction using 121 teachers from six schools in a single school district. For each of the four tasks performed by teachers, they found a

strong positive relationship between frequency of evaluation and satisfaction with the manner in which tasks were assigned and evaluated (Borucki & Scott, 1975).

When the same study was replicated using teachers in alternative schools, they found only a slightly positive correlation between frequency of evaluation and satisfaction (Borucki & Scott, 1975).

In another part of their research in the six schools, Borucki & Scott (1975) studied the differences in the perceptions of teachers and principals as to the frequency and content of evaluations. They found that for all four teaching tasks, principals believed themselves to be communicating evaluations more frequently than was reported by teachers. Principals also reported communicating negative evaluations far more often than teachers reported receiving them. Borucki and Scott (1974) considered this to be important because of the earlier research results in which they found a relationship between frequency of evaluation and satisfaction. They felt that because of the combined findings of the two studies, it is very important that a supervisor recognize the possibility of discrepant perceptions of frequency of evaluation, and particularly of communicated negative evaluation.

By comparing teachers in open-space classrooms with teachers in a traditional, one-teacher classroom, Borucki and Scott (1975) studied the relationship between the

visibility of work and a performer's satisfaction with evaluations. They found a strongly positive correlation between perceived visibility of work and satisfaction with evaluation. This concept of visibility is closely related to the sampling step in the six-stage model of the evaluation process.

In another study using data obtained from 131 teachers, the Borenbruch and Scott (1975) team analyzed the relationship between the power of subordinates and satisfaction with evaluation. The hypothesis that subordinates who believed they could influence the selection of criteria were more satisfied with the criteria and with evaluation in general was overwhelmingly supported.

Research in Universities

Using data collected in a large, private research university, Robert Hied (1971), under the supervision of Borenbruch and Scott, investigated authority relationships and evaluation processes in a university faculty. This research was the first attempt to study variations in authority system as a function of the type of task being performed. Hied (1971) selected four tasks which typically are performed by faculty: classroom teaching, research and scholarship, university service, and service external to the university.

Based on the idea that perceived disagreements among evaluators creates problems in authority structures, Lind theorized that a faculty member's satisfaction with the evaluation process is dependent upon the belief that the criteria are sufficiently clear so that all evaluators can agree on their evaluations. As expected, he found a strong relationship between perceived agreement among evaluators and satisfaction with the evaluation system. Although this relationship varied little from task to task, it did vary from discipline to discipline. Those disciplines believed by faculty to depend on a central body of theory were higher in perceived agreement among evaluators and higher in satisfaction with the evaluation system (Dornbusch & Smith, 1973). Lind (1971) reasoned that when evaluators do not agree on evaluations, faculty have more difficulty in reaching their personal goals through evaluations of their performance.

In another part of his research in universities, Lind assessed the relationship between the perceived amount of influence on rewards of a particular task and the amount of time spent on that task.

Among professors who saw teaching as Moderately, Very, or Extremely Influential, 45 percent were above the median in time spent on undergraduate teaching and 77 percent were above the median in time spent on graduate teaching. By contrast for those who viewed teaching as only Slightly or Not At All Influential, only 31 percent were above the median for undergraduate teaching and 47 percent were above the median for graduate teaching. Similarly, for research, those who saw

research as highly influential in the university's evaluation process were more likely to be above the median in time spent on research, compared with those who are research as less influential. The respective proportions being 43 percent and 48 percent. These data are persuasive in support of our hypothesized relationship since they demonstrate that the same persons who perceive either teaching or research to be more rewarded by the university are likely to spend more time on that activity. (Barnhouse & Scott, 1978, p. 387)

Based on additional data collected in this study, Lind concluded that since faculty expended considerable effort on teaching, they wanted a better balance between effort and rewards (Barnhouse & Scott, 1978).

Summarizing the vast decade-long research project, Barnhouse and Scott [1978] concluded:

The principle stated in this theory emphasizes the importance of evaluation processes. Performers and evaluators have their behavior shaped by evaluations made of them by others. Much that is right and much that is wrong in current organizations can be explained in terms of evaluation processes. The ostensible goals of organizations—their public goals—often are not built into the evaluation system. This lack results in goal displacement. We believe that evaluation is a fundamental process in all human interaction and specifically in the operation of authority systems in organizations. Accordingly, we must apply and extend knowledge of evaluation if we are to control the organizations that so often control us. (p. 387)

Contingency Theories

The contingency approach to leadership is one of the most popular approaches to the study of leadership and management styles. Contingency theories are based on the recognition that there is no one best way to lead and that what works well in one situation may not work well in

another. The basic assumption is that the proper match of leadership style and situation determines leadership success (Schriesheim, 1994). It is considered to be the middle ground between the view that there are universal principles of management and organization and the view that each organization is unique and that each situation must be analyzed separately (East & Rosenzweig, 1978).

East and Rosenzweig (1978) described contingency theory as follows:-

The contingency view seeks to understand the interrelationships within and among subsystems as well as between the organization and its environment and to define patterns or relationships or configurations of variables. It emphasizes the multivariate nature of organizations and attempts to understand how organizations operate under varying conditions and in specific circumstances. Contingency views are ultimately directed toward suggesting organizational designs and managerial actions most appropriate for specific situations. (p. 114)

The concept of task concepts (Bassburnh & Scott, 1978) is based on the contingency approach to leadership behavior. In this part of the literature review, the major approaches to contingency theory are reviewed, and the relationship between task concepts and contingency theories is established.

The first of the contingency theories was developed by Fred Fiedler (1977) after 30 years of research and gave its name from the following assumption:

The performance of a leader depends on two interrelated factors: (1) the degree to which the situation gives the leader control and influence—that is, the

Individual that he can successfully accomplish the job, and (2) the leader's basic motivation--that is, whether his self-esteem depends primarily on accomplishing the task or on having close supportive relations with others. (p. 29)

Fiedler believed that the degree of control was determined by three situational variables: (1) leader-member relations (good or poor)--the degree to which the group supports the leader; (2) task structure (high or low)--the extent to which task goals, procedures, and guidelines are clear and definable, and (3) position power (strong or weak)--the degree to which the position gives the leader the power to reward and punish subordinates. (Schermershoen, 1984).

Schermershoen (1984) summarized the results of Fiedler's research in two propositions:

Proposition No. 1 A task-oriented leader will be most successful in situations of either high or low control.

Proposition No. 2 A relationship-oriented leader will be most successful in situations of moderate control. (p. 319)

Fiedler's position was that effective leadership results from a proper match of leadership style and situation. Since he believed that leadership style is tied to basic personality characteristics and not easily changed, he advocated moving leaders to situations that matched their leadership style (Kouzes, 1998)

A relationship can be found between the task conceptualizations delineated by Borchert and Scott (1971) as varying

along a continuum from active to inert and the situational characteristics described by Fiedler as varying from low to control to high to control. Although Borchers and Scott focused on differences in tasks that an individual performs and Fiedler (1967) focused on the characteristics of different jobs within an organization, both share the commonality that different supervisory behaviors are effective in different situations. The task structure component used by Fiedler as part of his definition of low and high control is almost identical to the definition of task structure used by Borchers and Scott. It is difficult to directly compare the two theories, because Fiedler did not include evaluation as a component of his definitions of leader behavior, but similarities can be seen in the Fiedler conclusion that low control situations and high task-oriented leaders are most compatible and the Borchers and Scott finding that frequency of evaluation is more strongly related to satisfaction and coverage for tasks that are more active.

A second contingency approach to leadership--path-goal theory--has already been reviewed. Contrary to Fiedler's model, it is assumed in path-goal theory that the leadership style of an individual can vary in situations within the organization, vary due to personal characteristics of subordinates and due to environmental factors which stimulate, constrain, or reward the activation of workers. A leadership style will be effective as a path-goal means if it

satisfies needs and accommodates individual abilities and if it complements the degree of job structure that exists (Stogdins, 1971). Schriesheim (1986) listed some of the theory's basic predictions as

- Supportive leadership complements structured tasks.
- Directive leadership complements unstructured tasks.
- Achievement leadership complements achievement-oriented workers.
- Participative leadership complements highly capable workers. (p. 177)

These predictions are consistent with the Bennis and Scott (1975) and Stogdins (1974) concept of task conceptions. The path-goal postulate that directive leadership complements unstructured tasks might be considered an expansion for the Bennis and Scott and Stogdins concepts that frequency of activation is positively related to increased satisfaction and increased perceptions of effectiveness in relationship to effort. The idea that leaders can modify their leadership style to match the situation is supported by both theories.

Victor Vroom and Phillip Yetton (1975) offered another contingency theory that is based on how leaders make decisions. The central proposition in the Vroom/Yetton model is that effective leadership results when the decision method matches the characteristics of the problem.

Yessierli and Tettam (1975) identified these five decision-making styles, each requiring a different degree of subordinate participation:

1. Leader makes the decisions using whatever information is available.

2. Leader secures necessary information from members of the group, then makes the decision.

3. Leader shares the problem with relevant members of the group on a one-to-one basis, then makes the decision.

4. Leader shares the problem with members as a group at a meeting, then decides.

5. Leader, acting as a chairperson, shares the problem with the group and facilitates efforts of the group to reach consensus on a group decision.

By answering a series of questions, the proper decision-making style can be determined for each of 34 different decision situations. The questions relate to the necessity for a quality decision, the likelihood and necessity of acceptance by the subordinate, and agreement on organizational goals (Kreitner, 1988):

Although decision making and evaluation may not seem to be related, a review of the Bennis and Scott (1975) and Bennis (1984) six-step model of the evaluation process reveals a close interrelationship. The first three stages of the evaluation model--clarifying, criteria-setting, and sampling--involve significant decisions as to how work will

be done. If the seven Yess and Tenor questions (1973) are answered for different teaching tasks, different appropriate decision-making styles are indicated.

This is closely related to the Bennis and Bant (1975) concept of task conceptions in which they stated that subordinates desire different work and supervisory arrangements for different tasks. Bennis (1964) hypothesized that frequency of evaluation is more strongly related to increased perceptions of effectiveness in relationship to effort for tasks that are more active. This can be readily related to the Yess and Tenor proposition that structured and unstructured tasks require different levels of subordinate participation in the decision-making process.

Another contemporary theory of leadership that has attracted considerable attention is what Paul Hersey and Kenneth Blanchard (1977) called life-cycle theory or situational leadership theory. The theory is based on leadership studies known as the Ohio State Studies and on the work of William Reddin (Meyerson, Morley, & Piatra, 1988).

The basic concept of this theory is that a leader's behavior should be situational, based primarily on the maturity or immaturity of followers (Meyerson, Morley, & Piatra, 1988). Hersey and Blanchard (1977) defined maturity as the capacity of individuals to set high, realistic goals and as their willingness and ability to take responsibility for a specific task.

Bersy and Blanchard (1977) proposed that as the maturity level of subordinates increases, the most effective leadership style changes from high task-low relationship to moderate task/high relationship and, finally, to low task-low relationship. They defined task behavior as the extent to which leaders organize and define the roles of their followers and tell them when, where, and how tasks are to be accomplished. They defined relationship behavior as determined by the extent of support provided by the leader and the extent to which the leader engages in interpersonal communication and facilitating behavior (Hoganson, Mosley, & Plateri, 1988).

If faculty are considered to be highly mature individuals, a low relationship-low task leadership style is the most effective, according to the Bersy and Blanchard (1977) model. This seems to be directly contrary to the research by Barmuch and Davis (1973) and Sedwillo (1984) in which they found a positive relationship between frequency of evaluation and satisfaction and effectiveness in relationship to effort. Perhaps the two can be reconciled by considering the possibility that the pressures of declining enrollments, decreasing ability, less well-prepared students, and increased demands for accountability may have so weakened both the willingness and perceived ability of faculty that they can no longer be considered as at the

highest level of maturity, using the Hersey and Blanchard (1969) definition.

Summary

In this chapter, selected literature related to the effects of evaluation on faculty has been reviewed. The evolution of faculty evaluation has shown a movement from virtually no systematized evaluation prior to World War II to a tremendous interest in the mechanics of evaluation in the sixties and seventies to increased attention to the effects of evaluation on faculty in the eighties.

Although Bernbach and Scott (1975) in their research at Stanford were the first formally to study the relationship between evaluation and faculty satisfaction and to propose the idea of task conceptions as determinants of appropriate supervisory behavior, it is evident from the review of literature in this chapter that the constructs of the Bernbach and Scott theory are closely related to those of expectancy motivation theory, goal-goal theory, incentive motivational theory, and other contemporary theories.

The concept of leverage, developed by Katzellis (1984) as an extension of the Bernbach and Scott (1975) research, was found to be closely related to the expectancy portion of the formula in expectancy motivation theory in which a person's perception that their efforts lead to desired performance is held to be a determinant of motivation to perform a particular task. The concepts of expectancy (and

leveraged are further extended into the leadership domain by the path-goal theory postulate that leadership behavior is effective to the extent that it increases a subordinate's perception that their efforts will lead to desired performance and desired rewards. Thus, a theoretical basis for Maitlis's hypothesis that increased frequency of evaluation is related to increased perceptions of effectiveness in relationship to affect is provided by path-goal theory.

While there is disagreement among researchers as to the effects of evaluation on the intrinsic motivation of faculty, substantial evidence has been provided that evaluation may be an important factor in increasing the intrinsic satisfaction of faculty by increasing feelings of competence, responsibility, and control.

The Borsboom and Scott (1975) idea that different tasks require different supervisory behaviors to maximize effectiveness is supported by the contingency formulations of leadership. Borsboom and Scott went one step beyond the other contingency theories by breaking down the work of a single individual into different tasks requiring different leadership behaviors for maximum effectiveness.

In studies conducted in public schools, Borsboom and Scott (1975) and Maitlis (1988) have provided empirical evidence that frequency of evaluation is related to satisfaction and leverage. This relationship is firmly supported

by the several theories which have been reviewed in this chapter.

This project used constructs and techniques similar to those of Burnbach and Scott (1976) and Sotocailla (1984) to gather additional empirical evidence that this relationship is true for community college faculty.

CHAPTER 111

METHODS AND PROCEDURES

The Task

The purpose of this study was to examine the relationship between the perceptions of community college faculty of the frequency of evaluation activity and their perceptions of their own effectiveness in relationship to their effort on these six teaching tasks: classroom teaching, curriculum development, professional development, college service, student advising, and maintenance activities.

Specifically, the following questions were addressed:

1. For each of the six teaching tasks, was there a statistically significant association ($p < .05$) between perceived frequency of evaluation and perceived effectiveness in relationship to effort (leverage)?

2. Was there a stronger relationship between perceived frequency of evaluation and leverage for tasks that faculty perceived as active, for tasks faculty perceived as moderately active, or for tasks faculty perceived as inert?

3. For each of the six teaching tasks where there was a statistically significant association ($p < .05$) between perceived frequency of evaluation and leverage, was

perceived frequency of evaluation related to increases in perceived effectiveness, decreases in perceived effort, or both?

4. In which stages of the evaluation process did perceived frequency of evaluation have the strongest association with perceived effectiveness in relationship to effort [increases] for a combination of the six teaching tasks?

Questionnaire Design

A survey research design was adopted for this project, and a questionnaire was developed to determine the relationship between perceived frequency of evaluation and perceived effectiveness in relationship to effort for a variety of teaching tasks. The research design was similar to the design used by Katrielle (1964) in his study of the effects of evaluation on elementary school teachers. The questionnaire was designed to collect data in four areas:

1. Perceptions of respondents of the frequency with which they were evaluated in each of the five stages of the six-step model of the evaluation process (Katrielle, 1964) for six different teaching tasks. The fourth stage of the evaluation model--appraisal--was not included since this is not viable to a faculty member.

2. Perceptions of respondents of their effectiveness in relationship to effort for each of the six teaching tasks and for the combination of the six tasks.

2. Opinions of respondents of where they placed each of the six teaching tasks along the active-passive continuum described by Bornbaum and Scott (1975).

Questions related to these five steps of the evaluation model:

- Allocating
- Criteria Setting
- Sampling
- Communication
- Planning for Improvement

Questions related to these six teaching tasks:

- Classroom Teaching
- Curriculum Development
- Professional Development
- College Service
- Student Advising
- Maintenance Activities

The study was described in both the cover letter and the questionnaire as a study to investigate the working relationship between department chairmen (or other immediate supervisors) and faculty in community colleges and to investigate the effects of this relationship on the effectiveness of faculty. Each time the word evaluation was used, it was defined as any way a supervisor communicates either directly or indirectly how well the faculty member is doing on a particular task and includes everyday compliments, criticisms, suggestions, and nonverbal communications, as well as formal evaluation procedures. Examples of such

teaching task were included on the questionnaire (Appendix A).

The questionnaire consisted of five sections. Section I contained questions used to construct the measure of frequency of evaluation. Faculty were asked the following questions:

1. For each task, how often does your department chairman (or other immediate supervisor) tell you what should be accomplished (i.e., what your goals should be)?

2. For each task, how often do you learn of the criteria and standards that your department chairman (or other immediate supervisor) uses in supervising and evaluating your performance?

3. For each task, how frequently do you think your department chairman (or other immediate supervisor) observes aspects of your task performance?

4. How frequently do you receive feedback (formal or informal) on how well or poorly you are doing on each task?

5. How frequently does your department chairman (or other immediate supervisor) use the information from evaluations of your performance on these tasks to work together with you to plan ways to improve performance?

Faculty responded to the five questions for each of the six teaching tasks using Likert scale response categories. The scale was (1) Very Frequently, (2) Frequently, (3)

Sometimes, (4) seldom, and (5) almost never. The five questions corresponded to the five steps of the evaluation model.

Section II contained questions used to obtain two measures of leverage for each of the six teaching tasks. The first, the Faculty Assessment of Leverage (FAL), was determined by asking faculty this question:

Teaching tasks involve both some level of effort on your part and some level of payoff in terms of results. For each task, which of the following best describes the relationship between your effort and the payoff during this past year?

Faculty responded by selecting one of nine response categories composed of one of three ratings of effort (high, medium, low) and one of three ratings of payoff (high, medium, low).

A final question asked faculty to consider all of their teaching tasks together and to describe the relationship between their effort and the payoff in terms of results during the past year. The same response categories used for the FAL measurement for individual tasks were used.

The second measure of leverage, the Composite Assessment of Leverage (CAL), was constructed by asking faculty these two questions:

1. Compared to earlier in your faculty career, how effective have you been in achieving your goals for each task this year?

1. Compared to earlier in your faculty career, how hard are you working on each task this year?

Faculty responded for each of the six tasks using three response categories to describe their effortiveness: (1) Much Less Effective, (2) Less Effective, (3) About as Effective, (4) More Effective, and (5) Much More Effective. They responded using these response categories to describe their effort. (1) Not Nearly as Hard, (2) Not Quite as Hard, (3) About as Hard, (4) Somewhat Harder, and (5) Much Harder.

Section III contained questions designed to measure the task conceptions of faculty for each of the six teaching tasks as varying along the active-inert continuum described by Gornbeuch and Scott [1875]. These questions were asked for each of the six tasks to measure task conception in terms of clarity, predictability, and efficacy:

1. How vague or undefined do your goals for each task appear to you?

2. For each task, how often can you predict which way of doing things is most likely to enable you to reach your goals?

3. For each task, how often are you successful in reaching your goals?

The scale for the responses to question 1 was (1) Very Vague and Undefined, (2) Somewhat Vague and Undefined, (3) Occasionally Vague and Undefined, and (4) Almost Never Vague and Undefined. The scale for responses to questions 2

and 5 was (1) Very Frequently, (2) Frequently, (3) Sometimes, and (4) Almost never.

After the initial development of the questionnaire, it was administered to 25 community college faculty from three colleges that were not part of the sample. Each person was interviewed to determine which questions were unclear, misleading, or anxiety-producing. Modifications were made based on these recommendations. This process also provided the information that the average time to complete the questionnaire was 15 minutes.

Sample Selection

The target population for this study was all full-time faculty members in the 18 colleges in the Florida community college system. The list of faculty members for each college was based on the listing in the 1983-84 college catalog.

A multistage cluster sampling technique with probability proportional to size was used to limit the number of clusters in the sample to five, but, at the same time, to ensure that each faculty member at all colleges had an equal chance of selection. The first step in this process was to determine the total number of full-time faculty at each college. When a college had less than 100 faculty members, it was combined with the college nearest in geographic

location, and the resulting combination was treated as one cluster for sampling purposes.

Each cluster was assigned members in proportion to their size. For example, because Miami Dade had 184 of the total faculty, they were assigned 14 members and because Palm Beach had 4% of the total faculty, they were assigned 4 members. Five clusters were then selected using computer-generated random numbers.

After the five clusters were selected, 100 full-time faculty members were selected from each college using computer-generated random numbers. Each college was then asked to verify the list, and alterations were randomly selected as needed.

Data Collection Procedures

A cover letter (Appendix B) and the questionnaire (Appendix A) were sent to the 500 faculty members selected for the sample. A stamped return envelope was included. The participants were requested to complete the questionnaire and to return it in the enclosed return envelope.

A key administrator was selected at each college in the five clusters to assist in the follow-up of questionnaires that were not returned. A letter (Appendix C) was sent to each explaining the procedures and timetable for follow-up.

Three weeks after the original questionnaires were mailed, a list of non-responders was prepared for each

collapse. This was done from a code placed on the return envelope, because respondents were not asked to include their names on the questionnaire. This list was provided with additional questionnaires and stamped return envelopes to the project coordinator at each site. Each coordinator was asked to distribute these to non-responders and to encourage them to complete the questionnaire and to return it. At some institutions, this process had to be repeated a second time to meet the goal of an 80% response rate.

Analysis of the Data

To develop a measure of frequency of evaluation for each of the six teaching banks, respondents were divided into three groups. Those reporting no steps of the evaluation process in which they were evaluated either very frequently or frequently were classified as low frequency. Those reporting one step as being evaluated very frequently or frequently were classified as medium frequency, and those reporting two or more steps as being evaluated very frequently or frequently were classified as high frequency.

To obtain the RII measure of leverage for each of the teaching banks, faculty were classified according to those who chose a response category in Part 4 of Section II in which payoff was less than effort (leverage < 1) and those for whom payoff was equal to or greater than effort (leverage \geq or = to 1).

Responses to questions in Part 5 of Section II were used to calculate the second measure of leverage--the Composite Assessment of Leverage (CAL)--for each of the six teaching tasks. The CAL measure was constructed by taking the ratio of effectiveness to effort and placing faculty into two groups, those whose effectiveness was less than their effort and those whose effectiveness was equal to or greater than their effort.

To answer question number 1 of the problem statement, gamma describing the association between perceived frequency of evaluation and leverage were calculated for each of the six teaching tasks using both the TAL and CAL measure of leverage. Each gamma was tested for significance by transferring it into a normal distribution and calculating a z score using a two-tailed test and a 95% confidence level (Cron & Reed, 1980).

To answer question number 2 of the problem statement, each of the six teaching tasks was classified along the active-passive continuum using the responses to questions in Section III of the questionnaire. If a respondent selected either of the two most active response categories (very active/somewhat active for question 1, or sometimes/almost never for questions 2 and 3) for two or more of the three questions, the task was classified as active for that respondent. If a respondent selected either of the two most active response categories for one of the three questions,

the task was classified as moderately active for that respondent. If the respondent did not select either of the two most active response categories for any of the three questions, the task was classified as inert for that respondent.

The category which included the median for the respondents was used to classify each of the six tasks as active, moderately active, or inert. The data relating to perceived frequency and perceived leverage for tasks with a similar classification along the active-inert spectrum were collapsed. For example, since both classroom teaching and maintenance factors were classified as inert tasks, the data were added and the combination was treated as one task.

Gammas describing the relationship between perceived frequency of evaluation and perceived leverage were then calculated for the composite active, moderately active, and inert tasks. A test of significance was performed for each gamma, and the gammas were compared for strength of association to answer question number 1.

To answer question number 2 and to determine whether possible relationships between perceived frequency of evaluation and perceived leverage were due to decreases in perceived effort, increases in perceived effectiveness, or both, gammas were calculated using the separate effort and effectiveness components that were part of the DAL measure of leverage. One gamma was an expansion of the

relationship between perceived frequency of evaluation and perceived effectiveness, and another was an expression of the relationship between perceived frequency of evaluation and perceived effort. Both gamma's were tested for significance and compared for strength of association for each task where there was a significant association between perceived frequency of evaluation and perceived leverage.

To answer question number 4 concerning the strength of association between perceived frequency of evaluation in any single stage of the evaluation process and perceived leverage, gamma's were calculated for the combination of the six teaching tasks by classifying respondents into high, medium, and low categories of perceived frequency of evaluation for each of the five stages of the evaluation process. Responses were classified as high frequency for a given stage of the evaluation model if the respondent indicated very frequently or frequently for that stage for at least three of the six teaching tasks. Responses were classified as medium frequency if the respondent indicated very frequently or frequently for two of the six teaching tasks and were classified as low frequency if the respondent indicated very frequently or frequently for one or none of the six teaching tasks. The response to the last question in Section II, Part B, was used to determine leverage for the combination of all of the teaching tasks. The gamma's for each stage of the evaluation model were tested for

significance and compared with groups from the other stages for strength of association.

Summary

This chapter has presented the methods which were used to collect and analyze data to investigate the relationship between community college faculty perceptions of frequency of formal and informal evaluation and their perceptions of their own effort and effectiveness. Methods used to collect and analyze data relating to task conceptions as varying along an active-inert continuum and relating to the relative effects of evaluation in a single stage of the evaluation model were also described.

Since no assumption of distribution normality could be made, a nonparametric manner of analysis was used. Gamma, a non-parametric measure of association between two ordinal-level variables, was used to express the strength of the relationship between various combinations of perceived frequency of evaluation, perceived effort and effectiveness, and task conceptions. Each gamma was tested for statistical significance at the .05 level.

Chapter IV presents a description of the responses and the results of the analysis of these responses.

CHAPTER 18

RESULTS

This project was designed to determine whether there was a relationship between the perceived frequency of evaluation and perceived effort in relationship to effectiveness (leverage) of community college faculty. Ancillary purposes were to determine whether this relationship was stronger for tasks that were classified as more active along the active-inert continuum, whether there was a stronger relationship between perceived frequency and leverage in a particular stage of the evaluation process, and whether there was a stronger relationship between perceived frequency and perceived effort or between perceived frequency and perceived effectiveness.

The results of this project are presented in this chapter. The results are presented in three main sections: a description of the project sample and respondents, responses relating to perceived frequency of evaluation, responses relating to leverage, the relationship between frequency of evaluation and leverage, results of classifying tasks along the active-inert continuum, results from analyzing the effort and effectiveness components of leverage separately, and results from analyzing perceived

frequency of evaluation and perceived coverage in a single stage of the evaluation model for a combination of the six teaching tasks.

Description of Project Sample and Response

The results are based on 3 clusters that were selected from the 28 community colleges in Florida using a probability proportional to size sampling technique, with 100 faculty members selected from each of the 3 colleges.

The three clusters chosen for the sample represented colleges operating in urban and non-urban, union and non-union, and small and large environments. The results presented in Table 1 indicate that 81% of the total sample of 300 responded after multiple follow-up attempts, with a range of responses from individual colleges between 73% and 88%.

TABLE 1
COLLEGES INCLUDED IN THE SAMPLE AND NUMBER
AND PERCENTAGE OF RESPONSES

College	Number Sent	Number Returned	Percentage Returned
Everard	100	81	81%
Miami Dade	100	79	79%
Tallahassee	100	88	88%
Lake City			
North Florida	100	73	73%
Palm Beach	100	81	81%
Total	300	243	81%

Note: The 300 faculty surveyed represent 11.8% of the total population of community college faculty in Florida.

Responses Related to Perceived
Frequency of Evaluation

Perceived frequency of evaluation in each of the five stages of the evaluation model was calculated for each of the six teaching tasks. Responses indicating Frequently or Very Frequently were classified together as the highest frequency group, responses indicating Occasionally or Seldom were classified together as the middle group, and responses indicating Never were classified as the lowest group. These results are presented in Table 2.

TABLE 2
PERCEIVED FREQUENCY OF EVALUATION BY PROPORTION OF
RESPONDENTS FOR EACH STEP OF THE EVALUATION MODEL

Step	Frequency		
	Frequently or Very Frequently	Seldom or Occasionally	Never
Task: Classroom Teaching			
Allocating	.17	.39	.44
Criteria Setting	.23	.49	.28
Sampling	.14	.34	.52
Communicating	.18	.48	.34
Planning	.24	.34	.42
Average	.24	.45	.31
Task: Curriculum Development			
Allocating	.28	.34	.38
Criteria Setting	.23	.47	.30
Sampling	.20	.47	.33
Communicating	.29	.33	.38
Planning	.25	.42	.33
Average	.25	.44	.35

Table 2--Continued

Step	Frequency		
	Frequently or Very Frequently	Occasionally	Rarely
Task: Professional Development			
Allocating	.13	.46	.43
Criteria Setting	.38	.50	.38
Sampling	.13	.48	.37
Communicating	.38	.46	.36
Planning	.60	.40	.46
Average	.34	.48	.38
Task: Student Advising			
Allocating	.32	.46	.68
Criteria Setting	.38	.57	.54
Sampling	.34	.42	.66
Communicating	.66	.38	.53
Planning	.66	.42	.68
Average	.50	.48	.66
Task: College Service			
Allocating	.66	.34	.57
Criteria Setting	.66	.47	.57
Sampling	.67	.33	.56
Communicating	.66	.48	.57
Planning	.64	.42	.66
Average	.66	.48	.57
Task: No Inference Factors			
Allocating	.37	.47	.38
Criteria Setting	.38	.46	.37
Sampling	.37	.46	.37
Communicating	.38	.43	.36
Planning	.38	.44	.36
Average	.38	.44	.38

An analysis of the frequency distribution reveals that when responses for all tasks and all steps were combined, only 18% of the respondents indicated that they were evaluated either formally or informally on a frequent basis, and an average of 18% of the respondents selected the Never category for each task. Knowledge factors, a task defined as record keeping, office hours, and other routine activities, had the highest evaluation activity with 54% of the respondents reporting frequent evaluation activity for this task. Student advising had the lowest activity with only 14% of the respondents indicating frequent activity and 46% indicating no evaluation.

The step of the evaluation model which had the highest activity across all tasks was criteria setting, with an average of 54% of the respondents indicating frequent evaluation activity in this step. The steps indicating the lowest activity were allocating (goal setting) and planning for improvement with respondents indicating frequent activity only 17% of the time on each.

Measures Related to Leverage

Two measures of leverage were calculated from different items on the questionnaire. In questions relating to the first measure, the Faculty Assessment of Leverage (FAL), faculty were asked to describe their level of effort in relationship to the payoff in terms of results for each of

the two teaching tasks. If the respondent indicated that the payoff was greater than effort, the response was placed in the category indicating leverage greater than or equal to one. If the respondent indicated that payoff was less than effort, the response was placed in the category indicating leverage less than one.

On two questions relating to the Composite Assessment of Leverage (CAL), faculty described their present effort and effectiveness relative to their effort and effectiveness earlier in their careers. A ratio of the responses to the two questions was used to place the respondents in the leverage greater than or equal to one category or the leverage less than one category.

The results are presented in Table 3. The CAL measure was less variable than the PAL measure. At least 78% of all respondents reported leverage greater than or equal to one on the CAL measure. The lowest proportion of faculty were in the leverage greater than or equal to one category for the tasks of classroom teaching and maintenance factors using both the CAL and PAL measures. The highest proportion of faculty were in the leverage greater than or equal to one category for the task of student advising using both the CAL and PAL measures.

For the average of all tasks, 80% of the respondents indicated leverage greater than or equal to one using the PAL measure of leverage and 78% using the CAL measure. The

proportion for the CAL measure was consistently 15% to 33% larger than for the PAL measure. The range for the PAL measure was 51% to 73% and 73% to 83% for the CAL measure.

TABLE 3
PROPORTION OF RESPONDENTS INDICATING LEVERAGE
GREATER THAN OR EQUAL TO 1

Task	MEASURE OF LEVERAGE	
	PAL	CAL
Classroom Teaching	.52	.73
Curriculum Development	.46	.77
Professional Development	.48	.75
Student Advising	.72	.83
College Service	.55	.83
Maintenance Factors	.54	.75

The Relationship Between Perceived Frequency
of Evaluation and Leverage

To answer question number 1 of the problem statement and to determine whether there was a statistically significant association ($p < .05$) between perceived frequency of evaluation and perceived leverage for each of the six teaching tasks, respondents were categorized according to perceived frequency of evaluation and leverage using both the PAL and CAL measures of leverage. Respondents selecting either Frequently or Very Frequently for two or more stages of the evaluation scale were classified as High

frequency for this task, those selecting them for only one stage of the evaluation model were classified as medium frequency, and those selecting them in all stages were categorized as low frequency. The number of respondents in each category is reported in Table 4 for the FL measure of leverage and in Table 5 for the GAL measure of leverage.

TABLE 4
FREQUENCY OF EVALUATION AND LEVERAGE BY NUMBER OF
RESPONDENTS USING THE FL MEASURE OF LEVERAGE

Leverage	Frequency of Evaluation		
	Low	Medium	High
Task: Classroom Teaching			
Less than 1	98	48	59
Greater than or Equal to 1	138	82	33
Gamma = -.28 $\chi^2 = 2.98$			
Task: Curriculum Development			
Less than 1	188	26	38
Greater than or Equal to 1	184	24	87
Gamma = -.23 $\chi^2 = 1.81$			
Task: Professional Development			
Less than 1	128	31	38
Greater than or Equal to 1	148	52	44
Gamma = -.12 $\chi^2 = .891$			

Table 4-Continued

		Frequency of Satisfaction		
Leverage		Low	Medium	High
Task	Student Advising			
Less than 1		83	29	12
Greater than or Equal to 1		738	27	68
Sigma = 4.00		n = 841		
Task	College Service			
Less than 1		181	8	8
Greater than or Equal to 1		152	26	28
Sigma = 4.18		n = 3.051		
Task	Supervisory Factors			
Less than 1		189	8	31
Greater than or Equal to 1		129	22	22
Sigma = 4.06		n = .832		

TABLE 1

FREQUENCY OF EVALUATION AND LEVERAGE BY NUMBER OF
RESPONDENTS USING THE CAL MEASURE OF LEVERAGE

Leverage	Frequency of Evaluation		
	Low	Medium	High
Task: Classroom Teaching			
Less than 1	54	29	48
Greater than or Equal to 1	187	74	28
Gamma = -.38	$\chi^2 = 2.488$		
Task: Curriculum Development			
Less than 1	82	18	28
Greater than or Equal to 1	188	38	90
Gamma = +.38	$\chi^2 = 2.488$		
Task: Professional Development			
Less than 1	42	14	21
Greater than or Equal to 1	217	18	79
Gamma = +.88	$\chi^2 = .481$		
Task: Student Advising			
Less than 1	38	25	18
Greater than or Equal to 1	188	48	48
Gamma = +.18	$\chi^2 = 1.281$		

Table 3--Continued

Leverage	Frequency of Evaluation		
	Low	Medium	High
Task: College Service			
Less than 1	50	10	18
Greater than or Equal to 1	232	32	88
Gamma = .418	$\chi^2 = 2.188$		
Task: Maintenance Factors			
Less than 1	56	12	34
Greater than or Equal to 1	232	8	60
Gamma = -.31	$\chi^2 = .146$		

The values of gamma, a measure of strength of association between perceived frequency of evaluation and leverage are summarized in Table 4, for both the EAI and the FIA measures of leverage. Although the gammas using the FIA and EAI measures of leverage are different, the gammas using both measures are statistically significant at the .05 level for the same tasks (classroom teaching, curriculum development, and college service).

TABLE 4
CORREL FOR THE RELATIONSHIP BETWEEN PERCEIVED
FREQUENCY OF EVALUATION AND LEVERAGE

Teaching Task	MEASURES OF LEVERAGE	
	FIL	CAL
Classroom Teaching	-.23 ^a	-.34 ^a
Curriculum Development	+.33 ^a	+.34 ^a
Professional Development	+.11	+.09
Student Advising	+.09	+.11
College Service	+.25 ^a	+.36 ^a
Maintenance Factors	+.09	-.05

^a $p < .05$, two-tailed.

Using both the CAL and FIL measures of leverage, a significant and moderately negative relationship was found between perceived frequency of evaluation and leverage for classroom teaching. This indicated that as perceived frequency of evaluation increased, leverage decreased.

A significant and moderately positive relationship between perceived frequency of evaluation and leverage was found for the tasks of curriculum development and college service using both the CAL and FIL measures of leverage. This indicated that as perceived frequency of evaluation increased, leverage also increased.

No significant relationships were found at the .05 level of significance for the tasks of professional development, student advising, or maintenance factors using either the CRI or RAI measures of leverage.

Results from Classifying Tasks Along the Active-Inert Continuum

To determine where faculty placed each of the six tasks along an active-inert continuum, the questionnaire contained three questions requiring faculty to rate each task as to clarity, predictability, and efficacy. The task was classified as active for an individual respondent if the respondent selected one of the two most active response categories for two of the three questions, as moderately active if the respondent selected one of the two most active response categories for one of the three questions, and as inert if the respondent selected one of the two most active response categories for none of the three questions. The results of this classification are presented in Table 7.

Using data from Table 7, each task was classified as active, moderately active, or inert according to the category which included the median. The classification of each task is presented in table 8.

The data related to perceived frequency of evaluation and perceived leverage (RAI measure) were collapsed for tasks with the same classifications along the active-inert continuum, and new means were calculated for the combined

tasks and were tested for significance. Table 3 presents the data for the combined tasks.

TABLE 3
NUMBER OF RESPONDENTS CLASSIFYING EACH TASK AS
ACTIVE, MODERATELY ACTIVE, AND INERT

Task	Inert	Moderately Active	Active
Classroom Teaching	247	81	71
Curriculum Development	84	119	219
Professional Development	144	139	134
Student Advising	116	213	192
College Service	149	190	211
Maintenance Factors	127	71	111

TABLE 4
CLASSIFICATION OF TASKS ALONG THE
ACTIVE-INERT CONTINUUM

Task	Active-Inert Continuum
Teaching	Inert
Curriculum Development	Active
Professional Development	Moderately Active
Student Advising	Moderately Active
College Service	Active
Maintenance Factors	Inert

TABLE 4

DATA FOR THE RELATIONSHIP BETWEEN PERCEIVED FREQUENCY OF EVALUATION AND LEVERAGE FOR ACTIVE, MODERATELY ACTIVE, AND INERT TASKS

Task Category	Correl
Active (Curriculum Development and College Service)	.48 ^a
Moderately Active (Professional Development and Student Advising)	.28
Inert (Classroom Teaching and Maintenance Factors)	-.29 ^a

^a $p < .05$, two-tailed.

A significant and positive relationship between perceived frequency of evaluation and leverage was found for the combined active tasks. A significant and negative relationship was found for the combined inert tasks. A positive relationship was found for the moderately active tasks, but it was not significant at the .05 level.

Results from Analyzing the Effect and Effectiveness Components of Leverage Separately

The EAL measure of leverage was designed to permit an analysis of the effect and effectiveness components of leverage separately. Since leverage was defined as the ratio of effectiveness to effect, it would increase if

effectiveness increased as effort decreased. Table 10 presents an analysis of the relationship between perceived

TABLE 10
GAMMA FOR THE RELATIONSHIP BETWEEN PERCEIVED FREQUENCY
OF EVALUATION AND PERCEIVED EFFECTIVENESS
AND PERCEIVED EFFORT

Task	Gamma	
	Perceived Effectiveness	Perceived Effort
Tasks with Significant Association Between Perceived Frequency of Evaluation and Leverage		
Classroom Teaching	+.11	+.34 ^a
Curriculum Development	+.21 ^a	+.06
College Service	+.21 ^a	+.05
Tasks with No Significant Association Between Perceived Frequency of Evaluation and Leverage		
Student Advising	+.04	+.13
Professional Development	+.14	+.18
Maintenance Factors	+.02	+.14

^a $p < .05$

frequency of evaluation and perceived effectiveness and between perceived frequency of evaluation and perceived effort. The gammas were calculated by classifying responses

in the two highest response categories (More Effective/Much More Effective and Somewhat Harder/Much Harder) as high effectiveness or effort, by classifying responses in the middle response category (About as Effective or About as Hard) as medium effectiveness or effort, and by classifying responses in the lowest two response categories (Much Less Effective/Less Effective and Not Nearly As Hard/Not Quite as Hard) as low effort or effectiveness. Responses relating to frequency of evaluation were classified into high, medium, and low categories using the same techniques previously used in the study.

For the task of classroom teaching, although there was a slightly positive relationship between perceived frequency of evaluation and perceived effectiveness, there was a higher positive and significant relationship between perceived frequency of evaluation and perceived effort. Since leverage was defined as the ratio of perceived effectiveness to perceived effort, this significant positive association with effort is the reason for the significant negative association between perceived frequency of evaluation and perceived leverage for this task.

For the tasks of curriculum development and college service, a significant and uniformly positive relationship was found between perceived frequency of evaluation and perceived effectiveness. There was no significant relationship between perceived frequency of evaluation and

perceived effect for each task. For these tasks, the positive association between perceived frequency of evaluation and perceived effectiveness is the reason for the significant positive associations between perceived frequency of evaluation and perceived leverage.

Results from Analyzing Perceived Frequency
of Evaluation and Perceived Leverage in a
Single Stage of the Evaluation Model

Most of the analysis in this study involved investigating the relationships between perceived frequency of evaluation and leverage for each of the six teaching tasks using responses to all five steps in the evaluation model. However, an ancillary purpose of the project was to determine whether there was a stronger relationship between perceived frequency of evaluation in any single stage of the evaluation model and leverage for a combination of the six teaching tasks.

For each step of the evaluation model, an individual respondent was classified as high perceived frequency of evaluation for that step if the response was either Frequently or Very Frequently for three or more tasks, was classified as medium frequency if the response was either Frequently or Very Frequently for two tasks, and was classified as low frequency if the response was either Frequently or Very Frequently for zero or one of the six tasks. A composite measure of leverage was obtained from a question on

the survey instrument which asked respondents to describe their effort in relationship to the payoff (consulted) from that effort for all tasks combined. The distribution of responses for perceived frequency of evaluation and leverage in a single stage of the evaluation model is presented in Table 11.

TABLE 11
NUMBER OF RESPONDENTS BY PERCEIVED FREQUENCY OF
EVALUATION AND LEVERAGE FOR EACH STEP OF THE
EVALUATION MODEL

Step	Frequency of Evaluation		
	High	Median	Low
Allocating			
Leverage \geq or $=$ 1	71	8	112
Leverage $<$ 1	38	21	178
Options Setting			
Leverage \geq or $=$ 1	71	18	112
Leverage $<$ 1	48	56	148
Sampling			
Leverage \geq or $=$ 1	82	26	160
Leverage $<$ 1	28	20	172
Communicating			
Leverage \geq or $=$ 1	55	16	118
Leverage $<$ 1	88	28	148
Planning			
Leverage \geq or $=$ 1	71	8	112
Leverage $<$ 1	38	19	148

Kramer was calculated to determine the strength of association between perceived frequency of evaluation and

perceived leverage in a single stage of the evaluation model for a combination of the six tasks. These results are presented in Table 12.

TABLE 12
CORRELAS FOR PERCEIVED FREQUENCY OF EVALUATION AND
LEVERAGE IN A SINGLE STEP OF THE EVALUATION
MODEL FOR A COMBINATION OF ALL TASKS

Step	LEVERAGE
Allocating	+0.18 ^a
Criteria Setting	+0.12
Sampling	+0.14 ^a
Communicating	+0.13
Planning for Improvement	+0.42 ^a

^a $p < .05$, two-tailed.

Positive and significant associations between perceived frequency of evaluation and leverage for a combination of all tasks were found for the steps of allocating, sampling, and planning for improvement. There was no significant relationship between perceived frequency of evaluation and leverage for the steps of criteria setting and communicating.

A summary of these results and conclusions and recommendations drawn from these results are presented in Chapter 8.

CHAPTER V

SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

SUMMARY

Project Rationale and Design

Faced with declining enrollments, declining financial resources, an increasingly "homogenous," less-able, and aging faculty, and an increasing public outcry for accountability, community colleges have shown an increased interest in faculty evaluation as a possible solution to these problems. Most of this attention to evaluation has been focused on methods and procedures for evaluating faculty. Few researchers have concentrated on the effects of evaluation on faculty.

One of the first comprehensive studies was reported by Borchuck and Scott in 1978. They concluded that, based on studies conducted in schools, there was a relationship between frequency of evaluation and increases in teacher satisfaction. Borchuck and Scott (1978) also suggested that tasks vary along an active-passive continuum based on clarity, predictability, and efficacy and that performers' preferences for work arrangements and supervisory behavior vary according to their conceptions of various tasks. Borchuck and Scott also developed a four-step model

of the evaluation process which included the steps of allocating [task assignment and goal setting], criteria setting [establishing measurements], sampling [collecting information], and appraisal (performing the evaluation).

Wetzelite (1984) expanded the model to include the additional steps of communication and planning for improvement. Building on the work of Durebench and Scott (1975), he studied the effects of evaluation on the effort and effectiveness of elementary school teachers. He developed the concept of leverage to express the ratio between perceived effectiveness and perceived effort. He found a strong positive relationship between perceived frequency of evaluation and leverage and also found that perceived frequency of evaluation was related to increases in perceived effectiveness rather than to decreases in perceived effort.

To test the applicability of the Durebench and Scott (1975) and Wetzelite (1984) research to community colleges, this study was designed to determine the relationship between perceived frequency of formal and informal evaluation and perceived effectiveness in relationship to effort (leverage) for community college faculty. The main purpose of the study was to determine whether there was a statistically significant association at the .05 level between perceived frequency of evaluation and leverage for each of these six teaching tasks: classroom teaching, curriculum development, professional development, student

advising, college service, and maintenance factors. Auxiliary purposes of the project were to determine whether this association was stronger for tasks faculty perceived as active or for tasks they perceived as inert, whether there was a stronger relationship between perceived frequency of evaluation and perceived effectiveness or between perceived frequency of evaluation and perceived effort, and whether there was a stronger relationship between perceived frequency of evaluation and leverage in any one stage of the six-step evaluation model for a combination of all the teaching tasks.

A survey research design was adopted for this project, and a questionnaire was developed to determine perceived frequency of evaluation and perceived effectiveness and effort for each of the six teaching tasks and to determine the placement of each task along the activizer-inert continuum. The population was all full-time faculty in the 18 community colleges in Florida. A sample of five clusters of colleges was drawn using a probability proportionate to size sampling technique, and 100 faculty members were randomly selected from each of the five clusters. After multiple follow-up attempts, 82% of the sample of 100 responded to the mailing questionnaire.

The responses to the Likert-scale categories for each question were analyzed by constructing a summary measure of frequency of evaluation and two summary measures of

leverage (FMI and FMI) for each of the six tasks. Separate measures of perceived effectiveness and perceived effort were also constructed for each task. Using responses to questions related to clarity, predictability, and efficacy, each task was classified along an active-inert continuum, with active tasks defined as those tasks low on clarity, predictability, and efficacy. In addition, summary measures of frequency of evaluation in a single stage of the evaluation model for a combination of the six tasks and a measure of leverage for a combination of all tasks were constructed.

Gamma, a non-parametric measure of strength of association between ordinal-level variables, was used to describe the relationship between perceived frequency of evaluation and leverage, between perceived frequency of evaluation and perceived effectiveness, between perceived frequency of evaluation and perceived effort, and between perceived frequency of evaluation in each stage of the evaluation model and leverage for a combination of all tasks. In addition, data from similarly classified tasks along the active-inert continuum were collapsed, and gamma was calculated for each category so that relationships between perceived frequency of evaluation and leverage for active, moderately active, and inert tasks could be compared. Each gamma was tested for significance at the .05 level.

Results

A summary of the answers to each question in the problem statement is as follows:

Question 1. For each of the six teaching tasks, was there a statistically significant association ($p < .01$) between perceived frequency of evaluation and perceived effectiveness as relationship to effort (leverage)?

There was a statistically significant association ($p < .01$) between frequency of evaluation and leverage for the tasks of classroom teaching, curriculum development, and college service. There was no statistically significant association for the remaining three tasks.

Question 2. Was there a stronger relationship between perceived frequency of evaluation and leverage for tasks that faculty perceived as active, for tasks faculty perceived as moderately active, or for tasks faculty perceived as inert?

The strongest positive relationship was found for tasks perceived as the most active. However, a significant negative association was found for tasks perceived as inert.

Question 3. For each of the six teaching tasks where there was a statistically significant association ($p < .01$) between perceived frequency of evaluation and leverage, was perceived frequency of evaluation related to increases in perceived effectiveness, decreases in perceived effort, or both?

For the task of curriculum development and college service, there was a statistically significant association between perceived frequency of evaluation and perceived effectiveness and no statistically significant association between perceived frequency of evaluation and perceived effort. For the task of classroom teaching, there was a statistically significant association between perceived frequency of evaluation and perceived effort and no statistically significant association between perceived frequency of evaluation and perceived effectiveness.

Question 4 In which stages of the evaluation process did perceived frequency of evaluation have the strongest association with perceived effectiveness in relationship to effort (leverage) for a combination of the six teaching tasks?

The strongest significant relationships between perceived frequency of evaluation and leverage for a combination of the six teaching tasks were found in the stages of allocating, sampling, and planning for improvement. There was no significant relationship between perceived frequency of evaluation and leverage for the stages of criterion setting and communicating.

Analysis

Perceived frequency of evaluation and leverage No significant association was found at the .05 level between

perceived frequency of evaluation and leverage using either the GAI and PAI measures of leverage for the tasks of professional development, student advising, or maintenance factors. A significant, positive association was found for the tasks of curriculum development and college service, and a significant, negative association was found for the task of classroom teaching.

For the task of college service and curriculum development, there was a moderately positive relationship between perceived frequency of evaluation and leverage ($\gamma_{\text{GAI}} = .33$ using PAI and $.38$ using GAI). A similar relationship was found for the task of college service ($\gamma_{\text{GAI}} = .33$ using PAI and $.38$ using GAI). This would confirm the hypotheses of Bornbach and Scott (1975) and Katriella (1984).

For the task of classroom teaching, there was a moderately negative association between perceived frequency of evaluation and leverage ($\gamma_{\text{GAI}} = -.35$ using PAI and $-.38$ using GAI). These results are contrary to the hypotheses of Bornbach and Scott (1975) and Katriella (1984).

An analysis of these results indicated that there might be positive benefits in terms of perceived effectiveness in relationship to effort from increasing evaluation activity for the tasks of curriculum development and college service; in contrast, there might be a negative effect from increasing evaluation activity for the task of classroom

teaching. For the tasks of professional development, student advising, and maintenance factors, increases in perceived frequency of evaluation showed no significant relationship to leverage.

Classification of tasks along the active-passive continuum. Respondents identified curriculum development and college service as active tasks. Classroom teaching and maintenance factors were identified as inert tasks and professional development and student advising as moderately active tasks.

Since curriculum development and college service were the only two tasks identified as active tasks and were also the only two tasks where there was a significant, positive association of the .05 level between perceived frequency of evaluation and leverage, this would partially support the Matricilla (1984) hypothesis that there is a stronger relationship between perceived frequency of evaluation and leverage for tasks that are more active. However, Matricilla stated that this would not be true for the most active tasks since they were too active and unpredictable to be affected by evaluation activities. Curriculum development and college service were the most active tasks in this study so this would seem to disprove Matricilla's hypothesis. However, because he was unable to test this hypothesis because the teachers in his study perceived all of their tasks as inert and because there was not a clear definition of active

tasks reported in his research, it is possible that his intended definition and the one used in this study were different.

Support for the idea that task conceptualness affects the relationship between perceived frequency of evaluation and leverage was even stronger when the data for active tasks, moderately active tasks, and inert tasks were combined for the tasks in each category, and gamma was calculated for the relationship between perceived frequency of evaluation and leverage using the combined data. The gamma was .58 for active tasks, .28 for moderately active tasks, and -.18 for inert tasks. Although Katselias (1984) hypothesized that very inert tasks would show no relationship between perceived frequency of evaluation and leverage, in this study there was a significant, negative relationship between perceived frequency of evaluation and leverage for inert tasks.

The analysis of the data provided an affirmative answer to the problem statement which asked whether there was a stronger relationship between perceived frequency of evaluation and leverage for tasks that were more active than for tasks that were less active or inert.

Relationship between perceived frequency of evaluation and the separate effort and effectiveness components of leverage. The CM measure of leverage was designed to permit an analysis of the effort and effectiveness

components of leverage separately. The purpose of this analysis was to answer the question in the problem statement which asked whether there was a stronger association between perceived frequency of evaluation and perceived effectiveness or between perceived frequency of evaluation and perceived effort for tasks where there was a significant association at the .05 level between perceived frequency of evaluation and leverage.

There was no significant association between perceived frequency of evaluation and perceived effectiveness for the task of classroom teaching, but there was a significant and moderately positive association between perceived frequency of evaluation and perceived effort ($r_{\text{Pearson}} = .53$). This suggests that increases in the frequency of evaluation for this task were related to increases in effort but not to increases in effectiveness, thus explaining the negative association between perceived frequency of evaluation and leverage.

For the tasks of curriculum development and college service, there was a positive association between perceived frequency of evaluation and perceived effectiveness for both tasks ($r_{\text{Pearson}} = .41$ and $.42$, respectively) and no significant association at the .05 level between perceived frequency of evaluation and perceived effort. This indicated that increases in the frequency of evaluation were

related to increases in perceived effectiveness for these two tasks.

The conclusion was that for tasks where there was a significant negative association between perceived frequency of evaluation and perceived leverage, this was the result of a stronger association between perceived frequency of evaluation and perceived effort than between perceived frequency of evaluation and perceived effectiveness. For tasks where there was a significant positive association between perceived frequency of evaluation and perceived leverage, this was the result of a stronger association between perceived frequency of evaluation and perceived effectiveness than between perceived frequency of evaluation and perceived effort. This confirmed Bettelino's (1988) research which showed that increases in perceived frequency of evaluation were associated with increases in effectiveness, except that his data included no negative associations for comparison with the data from this study.

Perceived frequency of evaluation and leverage in a single stage of the evaluation model. When responses were analyzed for a single step of the evaluation model for a combination of all tasks, there were significant associations at the .45 level between perceived frequency of evaluation and leverage for the steps of allocating, sampling, and planning for improvement, with gamma of -.43, .39, and .41, respectively. There were no significant

associations at the .48 level between perceived frequency of evaluation and perceived leverage for the tasks of criteria setting and communicating.

The question in the problem statement which asked whether there was a stronger association between perceived frequency of evaluation and leverage in any single stage of the evaluation model was answered affirmatively. There was a stronger relationship between increases in perceived frequency of evaluation and leverage for the stages of eliciting, sampling, and planning for improvement than for the stages of criteria setting and communicating.

Conclusions and Implications

An analysis of the data collected in this study suggests that the leverage of community college faculty might be increased for certain tasks through more frequent evaluations of those tasks. This might be of great importance to community college administrators striving to find ways to overcome the inefficiencies of their faculty. Based on their data, those who perceive evaluation as a control mechanism and a threat to the intrinsic motivation and academic and professional freedom of faculty might consider the positive effects of increased evaluation on increasing faculty control over their own teaching tasks.

Although more research is needed to determine why there was a positive relationship between perceived frequency of

evaluation and leverage only for various tasks, an analysis of the data from this study tentatively suggests several possibilities. The tasks for which there was either a negative or no significant association between perceived frequency of evaluation and leverage, in addition to being less visible than the other tasks, were also tasks which were less visible to supervisors. Classroom teaching, student advising, and professional development, and in some cases maintenance factors are all very individual and private activities that often are not visible on a regular basis to supervisors. On the other hand, community college procedures usually require the frequent involvement of supervisors in curriculum development and college service activities, thus making the performance of faculty more visible for these tasks.

Another possible explanation for why there was a significant positive association between perceived frequency of evaluation and leverage only for the tasks of curriculum development and college service is that these two tasks might be the tasks for which faculty have the least training and least self-confidence and thus feel that they have the most to gain from evaluation activities. Also, supervisors might be considered by faculty to be more knowledgeable in these areas than in the other areas.

Perhaps the most dramatic and perplexing results of this study were the findings that faculty considered the

task of classroom teaching to be highly exact and that there was a significant negative association between perceived frequency of evaluation and leverage. When the wide variety of student abilities and backgrounds, the plethora of new research on topics such as learning styles and testing theory, and the increasing choices provided by developments in educational technology are considered, it is almost understandable that faculty consider classroom teaching to be high on predictability, clarity, and efficacy. Serious negative consequences, both for the work experiences of faculty and the learning experiences of students, might result from this finding.

Based on data collected in this study, the increased emphasis which is being placed on evaluation at many colleges might have negative effects on the perceptions of faculty of their own effort and effectiveness for the task of classroom teaching. This would create a situation where classroom teaching, the most important task of community college faculty, would be considered by faculty to be not only predictable and unchallenging but also to be a task where payoff in terms of results is less than effort. Community college administrators striving to improve classroom teaching might want to consider placing more emphasis on staff development, and less on staff evaluation, to assist faculty in gaining new skills or new perspectives to move classroom teaching back into the active task category.

Increased evaluation activity appears not to be the right solution to improving classroom teaching.

An analysis of the data also indicated that increasing certain types of evaluation may be more effective than other types in increasing leverage. There was no association between perceived frequency of evaluation and leverage for the steps of criteria setting and communicating, and there was a positive association for all other steps. It is possible that faculty responding to the questionnaire equated criteria setting and communicating to the formal college evaluation form and the required yearly supervisory evaluation. If this were true, these results might reflect a dissatisfaction with a process that is often considered to be too infrequent and arbitrary. The steps of discussing, sampling, and planning for improvement in the evaluation model might be used more frequently to increase faculty perceptions of effectiveness in relationship to effort.

Recommendations

Although much of the research of Rosenbach and Scott (1975) and Strickelle (1984) was supported by this study and additional data are now available with which to study the impact of evaluation processes on faculty, it is not a set of final answers. Additional data collected by other methods and in other settings are needed before conclusions can be reached about the role evaluation processes might

play in the operation of effective community colleges. The following are recommendations for further research.

1. The data collected in this study and the studies of Borchardt and Scott (1978) and Martocchio (1984) were based on teacher and faculty perceptions. Experimental confirmation of these findings is necessary if the direction of causation in the relationship between perceived frequency of evaluation and perceived effectiveness in relationship to effort is to be certain.

2. Other variables should be examined when determining which tasks show the strongest relationship between perceived frequency of evaluation and language. In particular, an analysis of the data collected in this study intuitively indicates that the visibility of task behavior may be a factor to be considered.

3. Data should be collected and analyzed from the same population using a personal interview technique to determine perceived frequency of evaluation and perceived effort and effectiveness. Data on frequency of evaluation, effort, and effectiveness should also be obtained from supervisors and compared with faculty data.

4. Organizational climate should be considered as a possible intervening variable in the relationship between perceived frequency of evaluation and language.

5. Research is needed to determine why classroom teaching is considered smart and to determine the implications of this task conception on teachers and on students.

6. The effects of supervisory training in performance appraisal and performance coaching techniques on the relationship between perceived frequency of evaluation and learning should be studied.

In conclusion, an analysis of the data from this research suggests that evaluation may affect the perceptions of community college faculty of their own effort and effectiveness. For some faculty the effect is positive and for others it is negative, and some types of evaluation activity have more effect than others. At the very least, it is a step towards silencing the opponents of faculty evaluation who have slowed progress in the study of evaluation, and it provides community college administrators with information that might be useful in designing both evaluation systems and faculty and supervisory development systems.

APPENDIX A
QUESTIONNAIRE

FACULTY TASKS AND SUPERVISORY RELATIONSHIPS

PURPOSE

The purpose of this study is to investigate the relationship between consistency in college faculty and department chairman or other immediate supervisors and to investigate the effects of this consistency on the effect of consistency on the performance of a variety of tasks. An hypothesis is that the study will result with results in favor of the one factor as a summary of the responses of the faculty members in the colleges.

SECTION ONE

DIRECTIONS

At each of the six teaching tasks described below, choose the five questions by circling an A in the space provided, indicating the frequency of the activity described in the question. One indicates that the activity occurs very frequently and five indicates that the activity occurs almost never.

When considering evaluation and supervisory relationship, it is important that you do not let any one of your supervisor's comments about this study or the study itself affect you in any way or in any way at all. This includes comments about the study and the study itself and responses to the study or to the study itself.

	1 Very Frequently	2 Frequently	3 Sometimes	4 Occasionally	5 Almost Never
Task: Classroom Teaching (Circle number indicating frequency of activity) 1. I give all personal opinions, conclusions and recommendations in my classroom (circle 1-5) 2. I give all personal opinions, conclusions and recommendations in my classroom (circle 1-5) 3. I give all personal opinions, conclusions and recommendations in my classroom (circle 1-5) 4. I give all personal opinions, conclusions and recommendations in my classroom (circle 1-5) 5. I give all personal opinions, conclusions and recommendations in my classroom (circle 1-5) 6. I give all personal opinions, conclusions and recommendations in my classroom (circle 1-5) 7. I give all personal opinions, conclusions and recommendations in my classroom (circle 1-5) 8. I give all personal opinions, conclusions and recommendations in my classroom (circle 1-5) 9. I give all personal opinions, conclusions and recommendations in my classroom (circle 1-5) 10. I give all personal opinions, conclusions and recommendations in my classroom (circle 1-5)					
Task: Classroom Teaching (Circle number indicating frequency of activity) 1. I give all personal opinions, conclusions and recommendations in my classroom (circle 1-5) 2. I give all personal opinions, conclusions and recommendations in my classroom (circle 1-5) 3. I give all personal opinions, conclusions and recommendations in my classroom (circle 1-5) 4. I give all personal opinions, conclusions and recommendations in my classroom (circle 1-5) 5. I give all personal opinions, conclusions and recommendations in my classroom (circle 1-5) 6. I give all personal opinions, conclusions and recommendations in my classroom (circle 1-5) 7. I give all personal opinions, conclusions and recommendations in my classroom (circle 1-5) 8. I give all personal opinions, conclusions and recommendations in my classroom (circle 1-5) 9. I give all personal opinions, conclusions and recommendations in my classroom (circle 1-5) 10. I give all personal opinions, conclusions and recommendations in my classroom (circle 1-5)					
Task: Classroom Teaching (Circle number indicating frequency of activity) 1. I give all personal opinions, conclusions and recommendations in my classroom (circle 1-5) 2. I give all personal opinions, conclusions and recommendations in my classroom (circle 1-5) 3. I give all personal opinions, conclusions and recommendations in my classroom (circle 1-5) 4. I give all personal opinions, conclusions and recommendations in my classroom (circle 1-5) 5. I give all personal opinions, conclusions and recommendations in my classroom (circle 1-5) 6. I give all personal opinions, conclusions and recommendations in my classroom (circle 1-5) 7. I give all personal opinions, conclusions and recommendations in my classroom (circle 1-5) 8. I give all personal opinions, conclusions and recommendations in my classroom (circle 1-5) 9. I give all personal opinions, conclusions and recommendations in my classroom (circle 1-5) 10. I give all personal opinions, conclusions and recommendations in my classroom (circle 1-5)					
Task: Classroom Teaching (Circle number indicating frequency of activity) 1. I give all personal opinions, conclusions and recommendations in my classroom (circle 1-5) 2. I give all personal opinions, conclusions and recommendations in my classroom (circle 1-5) 3. I give all personal opinions, conclusions and recommendations in my classroom (circle 1-5) 4. I give all personal opinions, conclusions and recommendations in my classroom (circle 1-5) 5. I give all personal opinions, conclusions and recommendations in my classroom (circle 1-5) 6. I give all personal opinions, conclusions and recommendations in my classroom (circle 1-5) 7. I give all personal opinions, conclusions and recommendations in my classroom (circle 1-5) 8. I give all personal opinions, conclusions and recommendations in my classroom (circle 1-5) 9. I give all personal opinions, conclusions and recommendations in my classroom (circle 1-5) 10. I give all personal opinions, conclusions and recommendations in my classroom (circle 1-5)					
Task: Classroom Teaching (Circle number indicating frequency of activity) 1. I give all personal opinions, conclusions and recommendations in my classroom (circle 1-5) 2. I give all personal opinions, conclusions and recommendations in my classroom (circle 1-5) 3. I give all personal opinions, conclusions and recommendations in my classroom (circle 1-5) 4. I give all personal opinions, conclusions and recommendations in my classroom (circle 1-5) 5. I give all personal opinions, conclusions and recommendations in my classroom (circle 1-5) 6. I give all personal opinions, conclusions and recommendations in my classroom (circle 1-5) 7. I give all personal opinions, conclusions and recommendations in my classroom (circle 1-5) 8. I give all personal opinions, conclusions and recommendations in my classroom (circle 1-5) 9. I give all personal opinions, conclusions and recommendations in my classroom (circle 1-5) 10. I give all personal opinions, conclusions and recommendations in my classroom (circle 1-5)					

QUESTIONS
PART B

For each of the six teaching topics, rate your current effort and effectiveness in comparison to that of a good teaching teacher. Place an X in the space provided for each. See Table below.

Remember to return to your rating period. How effectively have you rated it across your years for each topic listed on this page?

TEACHING TOPIC	Most Less Effective	Very Effective	Effective As Effective	Very Effective	Most More Effective
Classroom Management					
Classroom Organization					
Classroom Climate					
Classroom Management					
Classroom Management					
Classroom Management					
Classroom Management					

Remember to return to your rating period. How effectively have you rated it across your years for each topic listed on this page?

TEACHING TOPIC	Most Less Effective	Very Effective	Effectiveness As Effective	Very Effective	Most More Effective
Classroom Management					
Classroom Organization					
Classroom Climate					
Classroom Management					
Classroom Management					
Classroom Management					
Classroom Management					

TEACHING TOPICS
QUESTIONS

Remember to rate teaching topics by placing an X in the space to apply the rating. Remember to return to your rating period. Place an X in the space provided.

Remember to return to your rating period. How effectively have you rated it across your years for each topic listed on this page?

TEACHING TOPIC	Most Less Effective	Very Effective	Effectiveness As Effective	Very Effective	Most More Effective
Classroom Management					
Classroom Organization					
Classroom Climate					
Classroom Management					
Classroom Management					
Classroom Management					
Classroom Management					

2009/10/26/2010

Examine the six reading tasks by selecting the response category that best answers each question. Place an X in the space provided.

For each task, mark the response category that best answers the question. For each task, mark the response category that best answers the question.

Reading Task	Very Frequently	Frequently	Sometimes	Almost Never
Examine the Reading Task				
Examine the Reading Task				
Examine the Reading Task				
Examine the Reading Task				
Examine the Reading Task				
Examine the Reading Task				

For each task, mark the response category that best answers the question. For each task, mark the response category that best answers the question.

Reading Task	Very Frequently	Frequently	Sometimes	Almost Never
Examine the Reading Task				
Examine the Reading Task				
Examine the Reading Task				
Examine the Reading Task				
Examine the Reading Task				
Examine the Reading Task				

SECTION FOUR

Read the text and answer the questions (questions are listed below).

Read _____

Read _____

APPENDIX B

LETTERS

COVER LETTER SENT WITH QUESTIONNAIRE

Dear Colleagues:

I'm sure each of you can appreciate the excitement (and relief!) that I feel as I enter the final phases of my research for my doctoral degree in Educational Leadership. My research deals with the working relationships that community college faculty have with their supervisors and with how these relationships affect the effort and effectiveness of community college faculty.

As we search for ways to address the problems facing higher education in Florida, it is my hope that this research will provide data that will assist supervisors and faculty in maximizing their effort and effectiveness.

This questionnaire takes less than 15 minutes to complete. I know that your schedules are busy this time of year, but I'd really appreciate your assistance. A stamped return envelope is enclosed for your convenience. All replies will be strictly confidential. Data will be released only as a summary of responses from the faculty members in the five community colleges.

Thank you.

Sincerely,

Carel Cunningham

Telephone: (813) 341-3313 (Work)

LETTER SENT TO MY ADMINISTRATORS

Thank you for agreeing to assist me in following up the questionnaires sent to faculty at your college. I am enclosing additional questionnaires and a stamped return envelope for each faculty member from whom I have not received a reply. Except for the purposes of this follow-up, faculty were identified by number only and confidentiality was assured to each respondent. It would be best to ask them for their assistance in the study, but not to identify them as a non-respondent.

The questionnaire takes less than 15 minutes to complete, and the research might be best described as a study of the working relationships between department chairmen (or other immediate supervisors) and community college faculty and of the effects of these relationships on the effectiveness of faculty...

Again, thank you for your assistance. I'll be happy to share the results of the study with you when it is complete.

Sincerely,

Carol Capenbauer

REFERENCES

- Argyle, P. (1961). Interpersonal competence and attitudes. London: Allen Lane. Second ed., 1974.
- Bass, J. L. (1977). The motivation to teach. Journal of Higher Education, 48(3), 245-256.
- Bass, J. L. (Ed.). (1982). New directions for learning and teaching: Motivations involved in teaching. San Francisco: Jossey-Bass.
- Brown, J. A. (1975). Selection and evaluation of teachers. Berkeley, CA: McCutchan.
- Campbell, J., & Protherhard, R. (1958). Motivation theory in industrial and organizational psychology. In M. Brounstein (Ed.), Handbook of industrial and organizational psychology (pp. 261-293). Chicago: Rand McNally.
- Coxon, J. A. (1980). Determining faculty effectiveness. New York: Alfred A. Knopf.
- Chickering, A. (1984). Faculty evaluation: Problems and solutions. In P. Senneker (Ed.), Changing practices in faculty evaluation (pp. 51-84). San Francisco: Jossey-Bass.
- Cohen, A. M. (Ed.). (1976). New directions for counseling colleges, No. 1. Towards a professional faculty. San Francisco: Jossey-Bass.
- DeGroot, M. A. (1965). Criteria for measures of association. American Sociological Review, 30, 545-565.
- Ennis, P., Greenough, W. T., & Hanges, E. T. (1981). Student ratings of college teaching: Reliability, validity and usefulness. Review of Educational Research, 51, 415-435.
- Eriksonstehtelys, S. (1974). Intrinsic rewards and emergent motivation. In M. E. Lippert and S. Gerson (Eds.), The hidden roots of change (pp. 125-141). New York: Brunner.

- Chiksanterchelyi, M. (1982). Intrinsic motivation and affective learning: A flow analysis. In J. L. Bane (Ed.), New directions for teaching and learning: Motivational insights to teach affectively (pp. 11-19). San Francisco: Jossey-Bass.
- Cummings, L. L., & Schwab, D. P. (1975). Performance in organizations: Determinants and appraisal. Glenview, IL: Scott Foresman.
- Deci, E. L. (1975). Intrinsic motivation. New York: Plenum.
- Deci, E. L., & Ryan, R. M. (1982). The empirical exploration of intrinsic motivational processes. In L. Berkowitz (Ed.) Advances in experimental social psychology (pp. 111-145). New York: Academic Press.
- Deci, E. L., & Ryan, R. M. (1982). Intrinsic motivation by choice: Facilitation and obstacles in our colleges and universities. In J. L. Bane (Ed.), New directions for teaching and learning: Motivating teachers to teach affectively (pp. 27-35). San Francisco: Jossey-Bass.
- Bernbach, S. M., & Scott, R. W. (1975). Evaluation and the dynamics of leadership. San Francisco: Jossey-Bass.
- Boyle, R. D. (1981). Evaluating teaching. Lexington, MA: M. C. Heath.
- Bole, K. E. (1982). Professors as teachers. San Francisco: Jossey-Bass.
- Bole, K. E. (1978). The craft of teaching. San Francisco: Jossey-Bass.
- Boutin, M. E., & Ross, J. M. (1978). Increasing and decreasing intrinsic interest with contingent rewards: A test of cognitive evaluation theory. Journal of Experimental Social Psychology, 14, 488-500.
- Brans, M. (1978). Effects of supervisory behavior: Extensions of the path-goal theory of leadership. Journal of Applied Psychology, 63, 371-378.
- Fiedler, F. E. (1968). A theory of leadership effectiveness. 1968. New York: McGraw-Hill.
- Fiedler, F. E. (1977). Job engineering for effective leadership: A new approach. Management Science, 23(9), 10-18.

- Piller, A. G., House, R. M., & Kerr, S. (1971). Relationship of leader consideration and initiating structure to subordinate satisfaction. Administrative Science Quarterly, 16, 15-30.
- Salmonstich, J., & Cummings, J. L. (1967). An empirical investigation of the motivational determinants of task performance. Organizational Behavior and Human Performance, 1, 137-167.
- Georgopoulos, E. S., Mahoney, E. M., & Jones, M. W. (1957). A path-goal approach to productivity. Journal of Applied Psychology, 42, 349-355.
- Sass, G. (1966). The content of motivation. In P. Feldin (Ed.), Changing practices in family counseling (pp. 101-106). San Francisco: Jossey-Bass.
- Sass, G. (1968). Instrumentality theory of work motivation: Some experimental results and suggested modifications. Journal of Applied Experimental Research, 12, 1-15.
- Grandfield, L. B., & Neugarten, P. (1964). Supervisory characteristics and attitudes towards performance. Personnel Psychology, 17, 124-151.
- Hackman, J. B., & Wageman, E. B. (1995). Work redesign. Reading, MA: Addison-Wesley.
- Hagstron, W. O. (1965). The scientific community. New York: Basic.
- Hall, E. W., & Bergman, M. M. (1982). Organizational design and faculty motivation to teach. In J. L. Ross (Ed.), New directions for teaching and learning. Hillsdale, NJ: Jossey-Bass.
- Haneman, E. S., & Schuck, D. P. (1979). An examination of research on expectancy theory predictions of employee performance. Psychological Bulletin, 73, 1-6.
- Hersby, P., & Klenckow, E. (1977). Management of organizational behavior: Utilizing human resources (3rd ed.). Englewood Cliffs, NJ: Prentice-Hall.
- Hight, G. (1964). The art of teaching. New York: Alfred A. Knopf.

- Ward, S. P. (1971). Analysis of a faculty: Professionalism, evaluation and the authority structure. In J. E. Walbridge (Ed.), Academic Governance (pp. 181-215). Berkeley, CA: McCutchan.
- House, R. (1971). A path-goal theory of leader effectiveness. Administrative Science Quarterly, 16, 321-339.
- House, R. J., & Bennis, B. (1974). The path-goal theory of leadership: Some past, some and a priori tests. In J. G. Hunt (Ed.), Contemporary approaches to leadership (Carbondale, IL: Southern Illinois University Press).
- House, R., & Mitchell, T. R. (1974). Path-goal theory of leadership. Journal of Contemporary Business, 2, 81-97.
- Katz, P. E., & Brownstein, J. B. (1972). Contemporary ideas of administration and management. Chicago, IL: Science Research Associates.
- Kerr, G. (1971). Bureaucracy: Not as successful. In G. E. Smith (Ed.), New teachers, new schools. Current issues in higher education (pp. 7-12). San Francisco: Jossey-Bass.
- Kerr, G. (1974). Overcoming the dysfunctions of MBZ. Management by Objectives, 8(1), 74-87.
- Krogher, R. (1988). Management. Dallas: Houghton-Mifflin.
- Leish, J. A., & McInerney, W. J. (1975). The evaluation of teachers in higher education. In P. M. Werlinger (Ed.), Survey of research in education (Vol. 5, pp. 43-97). Ames, IA: Iowa.
- Loia, P., & Kauf, E. (1980). Basic statistics for managers. Boston: Little, Brown and Company.
- Lowler, E. (1971). Motivation in work organizations. Monterey, CA: Brooks/Cole.
- Lupper, R. R., & Gross, B. (1971). Turning play into work: Effects of adult surveillance and extrinsic rewards on children's intrinsic motivation. Journal of Personality and Social Psychology, 18, 119-127.
- Lewis, L. (1971). Scaling the career ladder. Baltimore, MD: The Johns Hopkins University Press.

- Lock, T. B. (1968). Merit rating: Have the issues changed? in Hopkinson: Learning professions (pp. 10-31). Washington, DC: American Association for Curriculum Development.
- Lockman, P., & Kruttsch, K. (1975). Pragmatism: A primer for applications. Glenview, IL: Scott, Foresman.
- McLaughlin, W. J. (1981). The rewards of teaching. In J. L. Ross (Ed.), New directions for teaching and learning: Motivating teachers to teach effectively (pp. 7-14). San Francisco: Jossey-Bass.
- Hopkinson, L. C., Ross, D. C., & Peters, P. (1968). Managerial concepts and applications. New York: McGraw-Hill.
- Miller, R. J. (1971). Evaluating faculty performance. San Francisco: Jossey-Bass.
- Miller, R. J. (1974). Developing programs for faculty evaluation. San Francisco: Jossey-Bass.
- Miller, J. A. (Ed.). (1983). Handbook of teacher evaluation. Beverly Hills, CA: Sage Publications.
- Nichol, C., Bernard, J. A., & Wilson, E. (1988). A test of expectancy motivation theory in educational organizations. Educational Administration Quarterly, 18(1), 79-92.
- Nichols, T. (1974). Expectancy models of job satisfaction, occupational preferences, and effort: A theoretical, methodological, and empirical appraisal. Psychological Bulletin, 81(1), 1983-1997.
- Morgan, M. M. (1994). Reward-induced decrements and increments in intrinsic motivation. Review of Educational Research, 64(1), 1-30.
- Mosley, R. T. (1983). Expectancy approaches to faculty motivation. In J. L. Ross (Ed.), New directions for teaching and learning: Motivating teachers to teach effectively (pp. 15-24). San Francisco: Jossey-Bass.
- Rotstein, G. (1984). Teachers' perceptions of the frequency of evaluation and measurements of their effort and effectiveness. Applied Behavioral Research Journal, 21(3), 379-385.

- Nie, N., Hull, C., Jenkins, J., Steinbrenner, K., & Bent, D. (1975). SPSS: Statistical package for the social sciences. New York: McGraw-Hill.
- Oberlander, D. (1984). A computer way to build faculty motivation (unpublished abstract, VI, 19). Austin: University of Texas, National Institute for Staff and Organizational Development.
- Poll, P. G., & Andrews, P. M. (1984). Encyclopedia in computer science. New York: John Wiley.
- Pross, R. J. (1983). Classroom: Cautions of learning laboratories. In C. A. T. Lee (Ed.), Improving college teaching (pp. 189-203). Washington, DC: American Council on Education.
- Ryan, R. M. (1982). Control and information in the leader-personnel system: An extension of cognitive evaluation theory. Journal of Personality and Social Psychology, 43, 458-481.
- Schwenkborn, J. A. (1984). Management for productivity. New York: John Wiley.
- Seedin, P. (1986). Successful faculty response programs. Englewood, CO: Corwin.
- Seldin, P. (1984). Changing practices in faculty roles. San Francisco: Jossey-Bass.
- Smith, M. B. (1976). The effects of social and cognitive rewards on intrinsic motivation. New York: Cornell University.
- Stra, B. (1978). Intrinsic and extrinsic motivation. Hackettstown, NJ: General Learning Press.
- Stein, R. M., & Porter, L. W. (1975). Motivation and work behavior. New York: McGraw-Hill.
- Sitanga, A. O., & Sims, M. P. (1978). An exploration of the path-goal theory of leadership in a health care environment. Academy of Management Journal, 21(4), 821-834.
- Thorn, Y. (1981). Task and motivation. New York: Wiley.
- Thorn, Y. M., & Thurn, P. M. (1975). Leadership and decision-making. Pittsburgh: Pittsburgh Press.

- Wilson, R. E. (1988). Establishing high commitment work systems. In J. E. Bishler, R. E. Wilson, & Associates (Eds.), The organizational life cycle (pp. 87-89). San Francisco: Jossey-Bass.
- Wardell, P. R. (1987). Contra and teacher evaluation. Washington, DC: United States Student Association.
- White, R. W. (1959). Motivation reconsidered: The concept of competence. Psychological Review, 66, 357-383.
- Wilson, L. (1942). The academic man. New York: Oxford University Press.
- Wilson, R. E. (1988). A case for merit pay. Community College Progress, 31(3), 28-30.
- Warrington, L. (Ed.). (1978). Remuneration opportunities for all: New staff for new schools. Washington, DC: American Association of Community and Junior Colleges.
- Zuckerman, M., Farns, J. P., Latham, G., Smith, K., & Reid, J. L. (1983). On the importance of self-determination for motivationally enhanced behavior. Personality and Social Psychology Bulletin, 9, 443-448.

BIOGRAPHICAL SKETCH

Carol Collins was born in Marion, Virginia, on March 21, 1948. She attended Marion High School, where she graduated as valedictorian of her class.

She attended the College of William and Mary in Williamsburg, Virginia, where she was the only female accounting major in the college. She was a member of Alpha Chi Omega Sorority and graduated from William and Mary with honors with a major in accounting and a minor in fine arts. She was selected as the Virginia Society of Certified Public Accountants Most Outstanding Senior Accounting Student in Virginia for 1968. She was awarded the degree of Master of Business Administration from the University of Florida in 1971.

Ms. Collins was a faculty member in the Division of Business at St. Petersburg Junior College from 1975 to 1981. From 1981 to 1983, she served as Division Secretary for Business Administration and Computer Science. In 1984, she was promoted to the position of Assistant Vice-President for Educational and Student Services. Her doctoral work in educational administration was at the University of Florida from 1984 to 1988.

Ms. Collins is the daughter of Ruth and Milton Collins of Naples, Florida, and the mother of a son, Brett Walker Copeland.

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.



David L. Anderson, Professor
Professor of Educational Leadership

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.



Dr. John W. Longfellow
Professor of Educational Leadership

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.



Dr. John E. Jones
Professor of Religion

This dissertation was submitted to the Graduate Faculty of the College of Education and to the Graduate School and was accepted as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

August 1988



Dean, College of Education

Dean, Graduate School